



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 171531

TO: Ginny Portner
Location: rem/3B02/3C18
Art Unit: 1645
Friday, December 02, 2005
Case Serial Number: 09/904994

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Portner,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library, Remsen Bldg.



STIC-Biotech/ChemLib

171531

MS

From: STIC-ILL
Sent: Monday, November 14, 2005 5:36 PM
To: STIC-Biotech/ChemLib
Subject: FW: 09/904,994

RECEIVED
NOV 15 2005
STIC-BIOTECH/CHEM LIB
(STIC)

-----Original Message-----

From: Portner, Ginny
Sent: Monday, November 14, 2005 5:29 PM
To: STIC-ILL
Subject: 09/904,994

Please search SEQ ID NO 1 with respect to Nucleic acid claims, and fragments thereof. thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

3C18

11/15/2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:29:19 ; Search time 14055 Seconds
(without alignments)
11659.884 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 rggagatttccacactt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: gb.hcg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2880.6	99.9	2883	6	BD185302	Helicobac
2	2880.6	99.9	2883	6	AX356683	Sequence
3	2228	77.3	2452	6	BD185306	Helicobac
4	2228	77.3	2452	6	AX356695	Sequence
5	2158.6	74.9	2405	6	BD185303	Helicobac
6	2158.6	74.9	2405	6	AX356686	Sequence
7	2158.6	74.9	2407	6	BD185305	Helicobac
8	2158.6	74.9	2407	6	AX356692	Sequence
9	1948	67.6	2183	6	BD185304	Helicobac
10	1948	67.6	2183	6	AX356689	Sequence
11	1185.6	41.1	8406	1	AF330621	Helicobac
12	1138	39.5	2664	1	HECUREASE	Helicobac
13	1090.8	37.8	12037	1	AE001446	Helicobac
14	1081.6	37.5	10903	1	AE000529	Helicobac
15	1075.2	37.3	2735	6	AR054309	Sequence
16	1075.2	37.3	5100	1	HECUREABCD	Helicobac
17	1072.4	37.2	7088	1	AB032429	Helicobac
18	1062.6	36.9	2619	6	A41006	Sequence 1

19	1062.6	36.9	2619	6	A72509	A72509 Sequence 1
20	1062.6	36.9	2619	6	A73285	A73285 Sequence 1
21	1062.6	36.9	2619	6	AR061045	Sequence
22	1062.6	36.9	2619	6	AR161678	Sequence
23	1060.4	36.8	2430	1	AY227442	Helicobac
24	1058.6	36.7	4824	6	AR349305	Sequence
25	1058.6	36.7	4824	6	AR349306	Sequence
26	1057.8	36.7	2746	1	HPFHUABG	X69080 H.felis ure
27	1056.8	36.7	2746	1	HPFHUABG	X17079 Helicobacte
28	1056.8	36.7	2767	6	A31515	A31515 H.pylori ur
29	985.2	34.2	6126	1	AF332656	Helicobac
30	985.2	34.2	308929	1	AE017145	Helicobac
31	894.6	31.0	1704	1	AY714224	Helicobac
32	885.8	30.7	1815	6	BD092530	Identific
33	884.8	30.7	1710	1	AY295085	Helicobac
34	881.8	30.6	1707	6	CQ903244	Sequence
35	881.8	30.6	1707	6	CQ903245	Sequence
36	880	30.5	1710	1	AF352376	Helicobac
37	878.4	30.5	1710	6	A07398	A07398 C.jejuni ge
38	878.4	30.5	1710	6	A08818	A08818 C.jejuni DN
39	876	30.4	1680	1	AF479026	Helicobac
40	772.4	26.8	110000	1	BA000004_02	Continuation (3 of
41	766	26.6	10118	1	BSUPEROP	Y08559 Bacillus su
42	766	26.6	200690	1	BSUB0019	Z39122 Bacillus su
43	739.4	25.6	1665	1	AF508003	AF508003 Helicobac
44	737.6	25.6	6537	1	BACUREA	D14439 Thermophili
45	733.8	25.5	4723	1	AB201709	AB201709 Campyloba

ALIGNMENTS

RESULT 1	BD185302	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302	Helicobacter felis vaccine.				
DEFINITION	BD185302	Helicobacter felis vaccine.				
ACCESSION	BD185302	GI:31877502				
VERSION	JP 2002355054-A/1.					
KEYWORDS	JP 2002355054-A/1.					
SOURCE	Helicobacter felis					
ORGANISM	Helicobacter felis					
REFERENCE	1 (bases 1 to 2883)					
AUTHORS	Kusters,J.G. and Cattoli,G.					
TITLE	Helicobacter felis vaccine					
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;					
COMMENT	AKZO NOBEL NV					
	OS Helicobacter felis					
	PN JP 2002355054-A/1					
	PD 10-DEC-2002					
	PF 16-JUL-2001 JP 2001214711					
	PI 17-JUL-2000 EP 00202565.8					
	PT JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI					
	PC					
	C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC					
	A61K39/23,A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15, PC					
	PC C12N1/19,					
	C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC					
	53,					
	PC					
	G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1.01),(C12Q1/68, PC					
	C12R1.01),					
	PC C12N15/00,C12N5/00,A61K37/02					
	CC Helicobacter felis vaccine					
	FT Key Location/Qualifiers					
	FT CDS (206)..(886)					
	FT CDS (897)..(2603).					
	Location/Qualifiers					
	1..2883					
	/organism="Helicobacter felis"					

/mol_type="genomic DNA" /db_xref="taxon:214"									
Query Match 99.9%; Score 2880.6; DB 6; Length 2883; Best Local Similarity 100.0%; Pred. No. 0; Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	RGRAGATTTCCARCACTTCAAGCACATATGATCTGTGTTGTGGGTGTAATTCR	60						
DB	1	RGRAGATTTCCARCACTTCAAGCACATATGATCTGTGTTGTGGGTGTAATTCR	60						
QY	61	ACTTGTAATRCATTAATTTTATAAATTAATTAATTAATTAATTAATTAATTA	120						
DB	61	ACTTGTAATRCATTAATTTTATAAATTAATTAATTAATTAATTAATTAATTA	120						
QY	121	TTACTTATATAAAGATTAATAAAGTAACGAAATTAAGACTATAATCCCATTCGCTT	180						
DB	121	TTACTTATATAAAGATTAATAAAGTAACGAAATTAAGACTATAATCCCATTCGCTT	180						
QY	181	TAAAAATTTAACACAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
DB	181	TAAAAATTTAACACAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
QY	241	GTATATATATCGGGCGAAGTGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCA	300						
DB	241	GTATATATATCGGGCGAAGTGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCA	300						
QY	301	ACCGAGACCATTTGCTTACATTAATGATGCCATATATGAGCAAGCGCGCTGGAAAAA	360						
DB	301	ACCGAGACCATTTGCTTACATTAATGATGCCATATATGAGCAAGCGCGCTGGAAAAA	360						
QY	361	AACCGTTGCCAGCTTATGGAAGATGTCATGCACTTTTGAAAAAGATGAAGTATGCC	420						
DB	361	AACCGTTGCCAGCTTATGGAAGATGTCATGCACTTTTGAAAAAGATGAAGTATGCC	420						
QY	421	CGGGGTGGTAATATGTTCCCGATCTAGGTGTAGAACCACTTTCTGATGTACGAA	480						
DB	421	CGGGGTGGTAATATGTTCCCGATCTAGGTGTAGAACCACTTTCTGATGTACGAA	480						
QY	481	ACTTGTAACCTGTGAATTTGGGCCATCGAACCCAGATGAGCACTTCAAAGCGGCGAAGTGAA	540						
DB	481	ACTTGTAACCTGTGAATTTGGGCCATCGAACCCAGATGAGCACTTCAAAGCGGCGAAGTGAA	540						
QY	541	ATTTGTTGCGATAAAGACATCGAGCTCAATGCAAGGCAAGAGCTTAACCGAATTTGAGGT	600						
DB	541	ATTTGTTGCGATAAAGACATCGAGCTCAATGCAAGGCAAGAGCTTAACCGAATTTGAGGT	600						
QY	601	TACTAATGAAGGGCTTAAATCCTTGCAATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAA	660						
DB	601	TACTAATGAAGGGCTTAAATCCTTGCAATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAA	660						
QY	661	CAAGGCACTAAATTCGATCGTGAAGCAAGCTATGGAAGCCCTAGATATTCCTCTGG	720						
DB	661	CAAGGCACTAAATTCGATCGTGAAGCAAGCTATGGAAGCCCTAGATATTCCTCTGG	720						
QY	721	CAACAGCTACGATTCGGGCGAGACAAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
DB	721	CAACAGCTACGATTCGGGCGAGACAAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
QY	781	CAGTAAAAAAGTATGGCATGAACGGCTTGTGAATAACATCGCGGATGAACGCCATAA	840						
DB	781	CAGTAAAAAAGTATGGCATGAACGGCTTGTGAATAACATCGCGGATGAACGCCATAA	840						
QY	841	ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900						
DB	841	ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900						
QY	901	AAATGAAAAACAGAATATGTAATACCTTACGACCCCAACCAAGGCGATTAAGTGCCT	960						
DB	901	AAATGAAAAACAGAATATGTAATACCTTACGACCCCAACCAAGGCGATTAAGTGCCT	960						
QY	961	TAGGAGATACCGATCTTTGGGCGAAGTAGAAACATGACTATATACCACCTATGGCGAAGAAC	1020						
DB	961	TAGGAGATACCGATCTTTGGGCGAAGTAGAAACATGACTATATACCACCTATGGCGAAGAAC	1020						
QY	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGCAAGTAATAGCCCTGATG	1080						
DB	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGCAAGTAATAGCCCTGATG	1080						
QY	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGCATGATTCGACTACACCGGATTTACA	1140						
DB	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGCATGATTCGACTACACCGGATTTACA	1140						
QY	1141	AGCCGCATTTGGATTTAAAAACGGCAAAATCCATGCAATTTGGCAAGCAAGAACAGG	1200						
DB	1141	AGCCGCATTTGGATTTAAAAACGGCAAAATCCATGCAATTTGGCAAGCAAGAACAGG	1200						
QY	1201	ACATGCAAGATGGCGTAAGCCCTCATATGGTCGTGGGTGTGGGCAACAGAAGCACTAGCAG	1260						
DB	1201	ACATGCAAGATGGCGTAAGCCCTCATATGGTCGTGGGTGTGGGCAACAGAAGCACTAGCAG	1260						
QY	1261	GGGAAGGTATGATTTATACCGCTGGGGAATCGAATCACACCCACCTTCCTTTCTCCAC	1320						
DB	1261	GGGAAGGTATGATTTATACCGCTGGGGAATCGAATCACACCCACCTTCCTTTCTCCAC	1320						
QY	1321	AACAAATTCCTTACCGCTTAGCCCAATGGCGTTACAAACCATGTTTGGAGGCGGCA	1380						
DB	1321	AACAAATTCCTTACCGCTTAGCCCAATGGCGTTACAAACCATGTTTGGAGGCGGCA	1380						
QY	1381	CTGTAGATGGCAGCAATGCGACTACTACTCTCGGGCAAAATGGAATTTGACCCGATGT	1440						
DB	1381	CTGTAGATGGCAGCAATGCGACTACTACTCTCGGGCAAAATGGAATTTGACCCGATGT	1440						
QY	1441	TGCGCGCAGCAGAAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
DB	1441	TGCGCGCAGCAGAAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
QY	1501	GCAAAAAACAACTTTGTAGAACAGTAGAAGCGGCGCATTTGGTTTAAATTTGCAATGAG	1560						
DB	1501	GCAAAAAACAACTTTGTAGAACAGTAGAAGCGGCGCATTTGGTTTAAATTTGCAATGAG	1560						
QY	1561	ACTGGGCGACAAACCAAGTCGATCGATCACTCTGAGCGTGGCAGATGAATAGCATG	1620						
DB	1561	ACTGGGCGACAAACCAAGTCGATCGATCACTCTGAGCGTGGCAGATGAATAGCATG	1620						
QY	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATAGGCGAGGTTATGTAGATGACCCCTAA	1680						
DB	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATAGGCGAGGTTATGTAGATGACCCCTAA	1680						
QY	1681	ATGCAATGAACGGGCGGCCATCCATGCTACCAATTTAGGAGCGGGTGGAGGACAT	1740						
DB	1681	ATGCAATGAACGGGCGGCCATCCATGCTACCAATTTAGGAGCGGGTGGAGGACAT	1740						
QY	1741	CACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
DB	1741	CACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
QY	1801	CTATTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGATGCCACC	1860						
DB	1801	CTATTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGATGCCACC	1860						
QY	1861	ACCTAGACAAACCGCATCCGAGGATTTACAAATTTCTCAAAGCGGTATCCGCCCGGCT	1920						
DB	1861	ACCTAGACAAACCGCATCCGAGGATTTACAAATTTCTCAAAGCGGTATCCGCCCGGCT	1920						
QY	1921	CTATCGGGCTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT	1980						
DB	1921	CTATCGGGCTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT	1980						
QY	1981	CGCAAGCAATGGGCGTGCAGCGAAGTATTCTCTCGAATTTGGCAGACTGCGGATAAGA	2040						
DB	1981	CGCAAGCAATGGGCGTGCAGCGAAGTATTCTCTCGAATTTGGCAGACTGCGGATAAGA	2040						
QY	2041	ATAAAAAAGAAATTTGGTAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGCAATTA	2100						
DB	2041	ATAAAAAAGAAATTTGGTAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGCAATTA	2100						

QY 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
DB |||||
QY 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
DB |||||
QY 541 ATTTGGTTCGATAAAGACATCGAGCTCAATGAGGCAAGAGTAAACCGAATCTGAGGT 600
DB |||||
QY 541 ATTTGGTTCGATAAAGACATCGAGCTCAATGAGGCAAGAGTAAACCGAATCTGAGGT 600
DB |||||
QY 601 TACTAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAA 660
DB |||||
QY 601 TACTAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAA 660
DB |||||
QY 661 CAAGGCATTAATAATTCGATCGTGAATAAGCGCTATGCGCAAAAGCGCTAGATATCCCTCTGG 720
DB |||||
QY 661 CAAGGCATTAATAATTCGATCGTGAATAAGCGCTATGCGCAAAAGCGCTAGATATCCCTCTGG 720
DB |||||
QY 721 CAACAGCTACGATTTGGGCGAGCAAAACCCGCAAGTGCAGTTCATCTCTGCTGGTGG 780
DB |||||
QY 721 CAACAGCTACGATTTGGGCGAGCAAAACCCGCAAGTGCAGTTCATCTCTGCTGGTGG 780
DB |||||
QY 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB |||||
QY 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB |||||
QY 841 ACATAAGCGCTTGACAAGCGGAAATCTCACGGATTTTCAAGTAAGGAGACTCCCATGA 900
DB |||||
QY 841 ACATAAGCGCTTGACAAGCGGAAATCTCACGGATTTTCAAGTAAGGAGACTCCCATGA 900
DB |||||
QY 901 AANATGAANAACAGATATGTAATACCTACGGACCCCAAGGCGATAAAGTGGCT 960
DB |||||
QY 901 AANATGAANAACAGATATGTAATACCTACGGACCCCAAGGCGATAAAGTGGCT 960
DB |||||
QY 961 TAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGACTATACCACTATGGGCGAAGAAC 1020
DB |||||
QY 961 TAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGACTATACCACTATGGGCGAAGAAC 1020
DB |||||
QY 1021 TTAATTTGGCGGGTAAACATATCCGTGAGGGTATGGGTCAAGCAATAGCCCTGATG 1080
DB |||||
QY 1021 TTAATTTGGCGGGTAAACATATCCGTGAGGGTATGGGTCAAGCAATAGCCCTGATG 1080
DB |||||
QY 1081 AAACACCCCTAGATTTAGTATCACTAAACCGGATGATTATCGACTACACCGGATTTTACA 1140
DB |||||
QY 1081 AAACACCCCTAGATTTAGTATCACTAAACCGGATGATTATCGACTACACCGGATTTTACA 1140
DB |||||
QY 1141 AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGAAACAGG 1200
DB |||||
QY 1141 AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGAAACAGG 1200
DB |||||
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAACACTAGCAG 1260
DB |||||
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAACACTAGCAG 1260
DB |||||
QY 1261 GGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACACACCACTTCTTTCTCCAC 1320
DB |||||
QY 1261 GGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACACACCACTTCTTTCTCCAC 1320
DB |||||
QY 1321 AACATTTCCCTACCGCTTAGCCCAATGGCGCTTACAACCATGTTTGGAGGGCGGACAGGTC 1380
DB |||||
QY 1321 AACATTTCCCTACCGCTTAGCCCAATGGCGCTTACAACCATGTTTGGAGGGCGGACAGGTC 1380
DB |||||
QY 1381 CTGTAGATGGCAAGATTCGACTACTACTCTCCGGGCAAAATGGAACTTGCACCGCATGT 1440
DB |||||
QY 1381 CTGTAGATGGCAAGATTCGACTACTACTCTCCGGGCAAAATGGAACTTGCACCGCATGT 1440
DB |||||
QY 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
DB |||||
QY 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
DB |||||
QY 1501 GCAAAAACACTTGTAGAACAGTAGAAGGGCGGATTTGGTATTTAAATTTGATGAAG 1560
DB |||||
QY 1501 GCAAAAACACTTGTAGAACAGTAGAAGGGCGGATTTGGTATTTAAATTTGATGAAG 1560
DB |||||
QY 1561 ACTGGGGCAACACCAAGTGGATCGATCACTCTTGGAGCGTGGCAGATGAATACGATG 1620

DB |||||
QY 1561 ACTGGGGCAACACCAAGTGGATCGATCACTGCTTGAAGCGTGGCAGATGAATACGATG 1620
DB |||||
QY 1621 TGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
DB |||||
QY 1621 TGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
DB |||||
QY 1681 ATGCAATGAACGGGCGGCGCATCCATGCTTACACATGAGGAGGCGGCTGAGAGACT 1740
DB |||||
QY 1681 ATGCAATGAACGGGCGGCGCATCCATGCTTACACATGAGGAGGCGGCTGAGAGACT 1740
DB |||||
QY 1741 CACTGATGTTATACCATTAATACCGTTCAGAGGAGCTCAATATTTCTACCCCTCTCCACACCCCA 1800
DB |||||
QY 1741 CACTGATGTTATACCATTAATACCGTTCAGAGGAGCTCAATATTTCTACCCCTCTCCACACCCCA 1800
DB |||||
QY 1801 CTATTTCCCTATACCATTAATACCGTTCAGAGGAGCTCAATGACATGCTCATGACATGCCACC 1860
DB |||||
QY 1801 CTATTTCCCTATACCATTAATACCGTTCAGAGGAGCTCAATGACATGCTCATGACATGCCACC 1860
DB |||||
QY 1861 ACCTAGACAAACCGCATCCGAGGATTTA CAATTTTCTCAAAGCGGATCCGCCCGGCT 1920
DB |||||
QY 1861 ACCTAGACAAACCGCATCCGAGGATTTA CAATTTTCTCAAAGCGGATCCGCCCGGCT 1920
DB |||||
QY 1921 CTATCGCGCTGAAGATGTCTCCATGATATGGGTGATCGCGATGACAGCTCGGAT 1980
DB |||||
QY 1921 CTATCGCGCTGAAGATGTCTCCATGATATGGGTGATCGCGATGACAGCTCGGAT 1980
DB |||||
QY 1981 CGCAAGCAATGGGCGGTGCGAGGAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
DB |||||
QY 1981 CGCAAGCAATGGGCGGTGCGAGGAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
DB |||||
QY 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTCGAAGATGCGAAAGATAACGATAAATTTCCGCAATTA 2100
DB |||||
QY 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTCGAAGATGCGAAAGATAACGATAAATTTCCGCAATTA 2100
DB |||||
QY 2101 AGCGCTACATCTCAAATACACTATCAACCCGCTTTGACCCACGGCGTGAGCGAGTATA 2160
DB |||||
QY 2101 AGCGCTACATCTCAAATACACTATCAACCCGCTTTGACCCACGGCGTGAGCGAGTATA 2160
DB |||||
QY 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGACTTTGGTGTGTGGAATCCTGCTTTTGTG 2220
DB |||||
QY 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGACTTTGGTGTGTGGAATCCTGCTTTTGTG 2220
DB |||||
QY 2221 GCGTAAAAACCAAAATCGTGAATCAAAGCGGTATGTTGTTCTCTGAAATGGCGAT 2280
DB |||||
QY 2221 GCGTAAAAACCAAAATCGTGAATCAAAGCGGTATGTTGTTCTCTGAAATGGCGAT 2280
DB |||||
QY 2281 CTAAACGGCTGTGCGCCACTCCCAACGGTTTATACCGGAAATGTTTGGGATCAG 2340
DB |||||
QY 2281 CTAAACGGCTGTGCGCCACTCCCAACGGTTTATACCGGAAATGTTTGGGATCAG 2340
DB |||||
QY 2341 GCAGGCGAAATTTGACACAGCATCACTTTTGTTCCAAAGTCCCTATGAAATGGG 2400
DB |||||
QY 2341 GCAGGCGAAATTTGACACAGCATCACTTTTGTTCCAAAGTCCCTATGAAATGGG 2400
DB |||||
QY 2401 TGAAGAAAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGGTAAACATCA 2460
DB |||||
QY 2401 TGAAGAAAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGGTAAACATCA 2460
DB |||||
QY 2461 CCAAGAAAGACTTCAAAGTTCAAACGCAAAAACGGCAAAAATCACCGTCAATCCGAAACCT 2520
DB |||||
QY 2461 CCAAGAAAGACTTCAAAGTTCAAACGCAAAAACGGCAAAAATCACCGTCAATCCGAAACCT 2520
DB |||||
QY 2521 TCGAGGCTTTGTAGATGGCAAACTCTGCACTCTTAACCCACCTCGCAAGTGCCTTAG 2580
DB |||||
QY 2521 TCGAGGCTTTGTAGATGGCAAACTCTGCACTCTTAACCCACCTCGCAAGTGCCTTAG 2580
DB |||||
QY 2581 CCCAGCGCTACACTTTCTTCTAGGCAAAATGCCCTTTGGGGCAGGTTATTTTAGGAA 2640
DB |||||
QY 2581 CCCAGCGCTACACTTTCTTCTAGGCAAAATGCCCTTTGGGGCAGGTTATTTTAGGAA 2640
DB |||||
QY 2641 TCTTTATCAAAACGCACTGCAATCGGTTCTGCTGTGCGATCGTGTGCTTTTAAACAAAC 2700
DB |||||

Db	2641	TTTTCATCAAAACGCACTCGTCAATCGGTCTTGGCTGTGCGATCGTGTGGCTTTAAAAACAAC	2700
Qy	2701	TTTTTCATCTTTAAAGCAATCGCCATTTTAAATTAATTTAAATTTCTTATAAATTAATTTATAT	2760
Db	2701	TTTTTCATCTTTAAAGCAATCGCCATTTTAAATTAATTTAAATTTCTTATAAATTAATTTATAT	2760
Qy	2761	TATGCCCTCTCATTTTTTAAAGAGAAATATATGCTAGGTCTTTGGTATATGCTATGTGGGG	2820
Db	2761	TATGCCCTCTCATTTTTTAAAGAGAAATATATGCTAGGTCTTTGGTATATGCTATGTGGGG	2820
Qy	2821	TTGTTTGGTGTGGGCGCAAGGGTATGAAACCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Db	2821	TTGTTTGGTGTGGGCGCAAGGGTATGAAACCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Qy	2881	AGG 2883	
Db	2881	AGG 2883	
RESULT 3			
BD185306			
LOCUS	BD185306	2452 bp	DNA linear PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.		
ACCESSION	BD185306		
VERSION	BD185306.1	GI:31877506	
KEYWORDS	JP 2002355054-A/5.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	1 (bases 1 to 2452)		
TITLE	Kusters, J.G. and Cattoli, G.		
JOURNAL	Helicobacter felis vaccine		
COMMENT	Patent: JP 2002355054-A 5 10-DEC-2002; AKZO NOBEL NV		
OS	Helicobacter felis		
PN	JP 2002355054-A/5		
PD	10-DEC-2002		
PP	16-JUL-2001 JP 2001214711		
PR	17-JUL-2000 EP 00202565.8		
PI	JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI		
PC	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC A61K39/23		
PC	A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15, PC C12N1/19,		
PC	C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC 53,		
PC	G01N33/53, G01N33/566, G01N33/569/(C12N9/80, C12P1:01), (C12Q1/68, PC C12P1:01),		
PC	C12N15/00, C12N5/00, A61K37/02		
CC	Helicobacter felis vaccine		
PH	Key	Location/Qualifiers	
FT	CDS	(48)..(728)	
FT	CDS	(739)..(2445).	
FEATURES	Location/Qualifiers		
source	1..2452		
ORIGIN	/organism="Helicobacter felis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:214"		
Query Match	77.3%;	Score 2228;	DB 6; Length 2452;
Best Local Similarity	94.3%;	Pred. No. 0;	
Matches 2312;	Conservative 0;	Mismatches 140;	Indels 0; Gaps 0;
Qy	159	AGGACTATAATCCCATTTGCTTTAAATTTAAACAAAGGAGTAATAGTGAACCTCACAC	218
Db	1	AGGACTATAATCCCATTTGCTTTAAATTTAAACAAAGGAGTAATAGTGAACCTCACAC	60
Qy	219	CCAAAGACGCAAGAAAGTTCTTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAG	278

Db	61	CCAAAGACGCAAGAAAGTTCTTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAG	120
Qy	279	CAGAGGGCTTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGTGCCTATATTATGG	338
Db	121	CAGAGGGCTTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGTGCCTATATTATGG	180
Qy	339	ACGAAGCGCGCCGTGGAAAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCATTTT	398
Db	181	ACGAGCGCGTCTGTGGCAAAAAAACCGTTGCCAGCTTATGGAAGAGTGTATGCATTTT	240
Qy	399	TGAAAAAGATGAAGTAATGCCCGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAAAG	458
Db	241	TGAAAAAGACGAGGTGATGCCCGGGTGGGGAATATGGTCCCTGATTTGGGCGGTGGAAAG	300
Qy	459	CCACCTTTTCTCTGATGGTACGAAACTTCTGTAATTTGGCCCATCGAACCCAGATGAGC	518
Db	301	CCACTTTTCCCGATGGCACCACAACTCTGTAATTTGGCCCATCGAACCTGATGAAAC	360
Qy	519	ACTTCAAAAGCGGGCGAAGTGAAATTTGGTTGCCATAAAGACATCGAGCTCAATGCAGGCA	578
Db	361	ACTTTAAGGGGGTGAAGTGAAATTTGGCTGTGATAAAGACATTGAACTCAACGCGGTA	420
Qy	579	AGAAAGTAACCGAACTTTGAGGTTACTAATGAAGGGCCTAAATCCTTGGCATGTGGGTAGCC	638
Db	421	AGGAAGTTACCGAACTAGAAGTTACTAAACGAAGGACCTAAATCCTTGCATGTGGGTAGCC	480
Qy	639	ATTTCACACTCTTTGAGCTAAACAGGCACCTAAATTCGATCGTGAAAGAGCCCTATGGCA	698
Db	481	ATTTCACACTCTTTGAGCCAAACAAAGCATTGAAATTCGATCGGGAAGAGCCCTATGGCA	540
Qy	699	AACGCTTAGATATTCCTCTGGCAACACGCTAGCAGTTGGGGCAGGCAAAACCCGCAAG	758
Db	541	AACGCTTAGATATTCCTCTGGCAACACACTACGATTTGGGCGAGCAAAACCCGTAAG	600
Qy	759	TGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGATTGGCATGAAAGGGCTTGTGAATA	818
Db	601	TGCAGTTAATTCCTCTTGGCGGTAGTAAAAAGTGATTGGCATGAAAGGGCTTGTGAATA	660
Qy	819	ACATCGCGATGAACCGCATAAACATAAAGCGCTTGACAGGCGAATCTCACGGATTTA	878
Db	661	ATATTGCGGACGAAACGCCATAAACACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA	720
Qy	879	TCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGATATGTAATACCTACGGACCC	938
Db	721	TCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGATATGTAATACCTACGGACCC	780
Qy	939	ACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC	998
Db	781	ACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC	840
Qy	999	TATACACCTATGGCGAAGAACTTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATG	1058
Db	841	TATACACCTATGGCGAAGAACTTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATG	900
Qy	1059	GGTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTGTATCATCTACACGGCATGATT	1118
Db	901	GGTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTGTATCATCTACACGGCATGATT	960
Qy	1119	ATCGACTACACCGGGATTTACAAAGCGGACATTTGGGATTTAAAAACGGCAAAATCCATGGC	1178
Db	961	ATTGACTACACCGGGATTTACAAAGCGGACATTTGGCATTAAAAATGGCAAAATCCATGGC	1020
Qy	1179	ATTGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGGT	1238
Db	1021	ATTGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGGT	1080
Qy	1239	GTGGGCACAGAAGCACTAGCAGGGGAGGTATGATTTATTTACCGCTGGGGGAACTCGATTCA	1298
Db	1081	GTGGGCACAGAAGCACTAGCAGGGGAGGTATGATTTATTTACCGCTGGGGGAACTCGATTCA	1140
Qy	1299	CACACCACCTCTCTCTTCTCCACAAATTCCTTACCGCTCTAGCCAAATGGCGTTTACAACC	1358

```
Db 1141 CACACCCACTTCTCTCTCCACAAACAAATTCCCTTACCGCTTAGCCAAATGCGGTTCACAA 1200
Qy 1359 ATGTTTGGAGCGGCA CAGGTCCTGTAGATGGCAAGATTCGCACTACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCA CAGGCCCGTAGATGGCAAGATTCGCACTACTACTATCACTCCGGGC 1260
Qy 1419 AAATGGAACCTTGACCGCATGTTGCGGCGAGAGAGATTTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACCTTGACCGCATGTTGCGGCGAGAGAGATTTCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGGCAAGGCAATAGCTTAGCAAAAAACAACTTGTGAACAAGTAGAAGCGGGCGG 1538
Db 1321 TTGGGCAAGGCAATAGCTTAGTAAAAACAACCTTGTGAACAAGTAGAAGCGGGCGG 1380
Qy 1539 ATTGGTTTTAAATPGCATGAAGACTGGGGCACAAACCAAGTCGATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTTAAATPGCATGAAGACTGGGGCACAACTCCAAGTCGATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACACCGATACGTTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCCCTTAATGCAATGAACGGGCGGCCCATCCATGCTTACCACATT 1718
Db 1501 GGTATGTAGATGACACCCCTTAATGCAATGAACGGGCGGCCCATCCATGCTTACCACATT 1560
Qy 1719 GAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCGAGTGAATATT 1620
Qy 1779 CTACCTCTCTCCACACCCCTACTATCCCTATACCAATTAATACGGTTGCAGAACCTTA 1838
Db 1621 CTACCTCTCTCCACACCCCTACTATCCCTATACCAATTAATACGGTTGCAGAACCTTA 1680
Qy 1839 GACATGCTATGATGACGACCCACCTAGACAAAGCATCCGAGGAGATTTACAATTTTCT 1898
Db 1681 GACATGCTATGATGACGACCCACCTAGACAAAGCATCCGAGGAGATTTCCAATTTTCT 1740
Qy 1899 CAAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTGCTCCATGATATCGGTGTG 1800
Qy 1959 ATCCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTGA 2018
Db 1801 ATCCGATGACAAGTTCGGATTCGCAAGCAATGGGCGCTGGGGAAGTGAATTCCTAGA 1860
Qy 2019 ACTTGGCAGACTGGGATAGAAATAAAAAAATTTGGTAAGCTTCTCTGAAGATGGCAAA 2078
Db 1861 ACTTGGCAAACTGCAGACAAGAAATAAAAAAAGATTTGGTAAGCTTCTCTGAAGATGGTGA 1920
Qy 2079 GATAACGATTAATTTCCGCATTAAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAATTTCCGCATCAAAACGCTATATCTCCAAATACACTTAATTCCTCGCTTTG 1980
Qy 2139 ACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTG 2198
Db 1981 ACCATCGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTG 2040
Qy 2199 GTGTGGAATCTGCTTTTGGCGTAAAAACCCAAAATCGTGATCAAGCGGATGATGGTG 2258
Db 2041 GTGTGGAATCTGCTTTTGGCGTAAAAACCCAAAATCGTGATCAAGGCGGATGATGGTG 2100
Qy 2259 GTCTTCTCTGAAATGGGCGATTTCTAACGCGTCTGTGCCACTCCCGCAACCGGTTTATTAC 2318
Db 2101 GTGTCTCTGAAATGGGCGATTTCTAATGCGTCTGTGCCACTCTCTCAGCCGGTTATTATAC 2160
Qy 2319 CGCGAAATGTTTGGGCGATCAGCGCAAGCGCAAAATTTGACACAGCATCACTTTTGTTCCT 2378
Db 2161 CGCGAAATGTTTGGGCGATCAGCGCAAGCGCAAAATTTGACACAGCATCACTTTTGTTCCT 2220
Qy 2379 AAAGTCGCTATGAAATATGGCGTGAAGAAAGCTGGGCTTAGACGCGCAAGTCTACCG 2438
Db 2221 AAAGTCGCTATGAAATATGGTGTGAAGAAAGAACTAGGTTTAGAGCGCAAGGTGCTCCCC 2280
```

```
Qy 2439 GTCAAAAACTGCCGTAAATCATCAACAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAA 2498
Db 2281 GTCAAAAACTGCCGTAAATCATCAACAGAGAGACTTCAAGTTCAACGACAAACCTGCAAAA 2340
Qy 2499 ATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGCGAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGCGAAACTCTGCACCTCTAAA 2400
Qy 2559 CCCACCTCGCAAGTGCCTCTAGCCACGCTACACTTTCTTAGGCACAAT 2610
Db 2401 CCCACCTCGAAGTGCCTCTAGCCACGCTACACTTTCTTAGGCATAAT 2452
```

```
RESULT 4
AX356695
LOCUS AX356695 2452 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 13 from Patent EP1176192.
ACCESSION AX356695
VERSION AX356695.1 GI:18674032
KEYWORDS
SOURCE
ORGANISM
Helicobacter felis
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
1
```

```
REFERENCE
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 13 30-JAN-2002;
Akzo Nobel N.V. (NL)
```

```
FEATURES
source
1. 2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPEQKFLYYAGEVARKKAEGLKLNQPEALIVISAHIM
DEARRKKTVAEIMEECHEFLKDEVMGVMVDPDLEATFPDGTGKILGIGKAGNKD
DEHFKAGEVFGCDKDIELNAGKEVTELEVNEGFKSLHVGSHFFPANKALFKFRE
KAYGKRLDIPSGNTLRICAGQTRKQLIPLGSSKKVIGMNLVNNIADERHKKALDK
AKSHGPTK"
739..2445
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQRYVNTYGTPTGDKVRLGDTDLWAEVEHDYTTTVEELKF
GAGKTIREGMGQSNPDNTLDLVITNMIIDYGYKADIGIKNGKILGIGKAGNKD
MQGVSHPMVVVGTEALAGGEMIITAGGIDSHTHFLSPQFPFALANGVITMFGGTT
GPVDGTNATITPGKWNLMRLRAEYSMMVFLGSSSSKKQLVQVEAGATGFK
LHEDGTTGPAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLANMNGRAIHAYHIEG
AGGCHSPDVIIMAGEVNTLPSSTTPTIPTINTVAEHLDMLTCHHLDKRIREDLPFS
QSRIRPSIAEDVLHIDGVIATMTSSDSQAMGRAGEVIPRTMQTADKNKKEKLPED
GAMDNFRIKRYISNTINPALTHGVSEYIGSVEEGKIADLVNWNPAFGVAPKPIVIK
LVVMVFMSGDSNASVPTPOPYREMFHGHGAKFDTSITFVSKVAYENGVKELGL
ERKVLVKNCRNITKQDFKENDKAKITVDPKTFEVDGKLCSTSKPTSEVPLAQRYT
FF"
```

```
CDS
1. 2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPEQKFLYYAGEVARKKAEGLKLNQPEALIVISAHIM
DEARRKKTVAEIMEECHEFLKDEVMGVMVDPDLEATFPDGTGKILGIGKAGNKD
DEHFKAGEVFGCDKDIELNAGKEVTELEVNEGFKSLHVGSHFFPANKALFKFRE
KAYGKRLDIPSGNTLRICAGQTRKQLIPLGSSKKVIGMNLVNNIADERHKKALDK
AKSHGPTK"
739..2445
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQRYVNTYGTPTGDKVRLGDTDLWAEVEHDYTTTVEELKF
GAGKTIREGMGQSNPDNTLDLVITNMIIDYGYKADIGIKNGKILGIGKAGNKD
MQGVSHPMVVVGTEALAGGEMIITAGGIDSHTHFLSPQFPFALANGVITMFGGTT
GPVDGTNATITPGKWNLMRLRAEYSMMVFLGSSSSKKQLVQVEAGATGFK
LHEDGTTGPAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLANMNGRAIHAYHIEG
AGGCHSPDVIIMAGEVNTLPSSTTPTIPTINTVAEHLDMLTCHHLDKRIREDLPFS
QSRIRPSIAEDVLHIDGVIATMTSSDSQAMGRAGEVIPRTMQTADKNKKEKLPED
GAMDNFRIKRYISNTINPALTHGVSEYIGSVEEGKIADLVNWNPAFGVAPKPIVIK
LVVMVFMSGDSNASVPTPOPYREMFHGHGAKFDTSITFVSKVAYENGVKELGL
ERKVLVKNCRNITKQDFKENDKAKITVDPKTFEVDGKLCSTSKPTSEVPLAQRYT
FF"
```

```
CDS
739..2445
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQRYVNTYGTPTGDKVRLGDTDLWAEVEHDYTTTVEELKF
GAGKTIREGMGQSNPDNTLDLVITNMIIDYGYKADIGIKNGKILGIGKAGNKD
MQGVSHPMVVVGTEALAGGEMIITAGGIDSHTHFLSPQFPFALANGVITMFGGTT
GPVDGTNATITPGKWNLMRLRAEYSMMVFLGSSSSKKQLVQVEAGATGFK
LHEDGTTGPAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLANMNGRAIHAYHIEG
AGGCHSPDVIIMAGEVNTLPSSTTPTIPTINTVAEHLDMLTCHHLDKRIREDLPFS
QSRIRPSIAEDVLHIDGVIATMTSSDSQAMGRAGEVIPRTMQTADKNKKEKLPED
GAMDNFRIKRYISNTINPALTHGVSEYIGSVEEGKIADLVNWNPAFGVAPKPIVIK
LVVMVFMSGDSNASVPTPOPYREMFHGHGAKFDTSITFVSKVAYENGVKELGL
ERKVLVKNCRNITKQDFKENDKAKITVDPKTFEVDGKLCSTSKPTSEVPLAQRYT
FF"
```

```
ORIGIN
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCCATTCGCTTTAAATTTAAACAAGAGGTAATAGGTGAACACTCACAC 218
Db 1 AGGACTATAATCCCATTCGCTTTAAATTTAAACAAGAGGTAATAGGTGAACACTCACAC 60
```

QY 219 CCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCTAGAAAGCGCAAG 278
DB 61 CCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCCATATTATGG 338
DB 121 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCCATATTATGG 180
QY 339 ACGAAGCGGCGCGTGGAAACCAACCGTTGCCAGCTTATGGAAGAGTGCAATGCATTTT 398
DB 181 ACGAGCGCGTCTGGCAACCAACCGTTGCCAGCTTATGGAAGAGTGATGCATTTT 240
QY 399 TGAACCAAGATGAAGTAATGCCCGGGGTGGTAAATATGTTCCGATCTAGAGTGTAAG 458
DB 241 TGAACCAAGATGAAGTGATGCCCGGGGTGGGAAATATGTTCCGATCTAGAGTGTAAG 300
QY 459 CCACCTTTCTGATGTGTAACACTGTAACTGTGAATTTGGCCCATCGAACCAAGATGAGC 518
DB 301 CCACCTTTCCCGATGGCACCAACTCGTAACCTGTGAATTTGGCCCATCGAACCTGATGAAC 360
QY 519 ACTTCAAAGCGGGCGAAGTGAAATTTGGTTGCGATAAAGACATCGAGCTCAATGCGAGGCA 578
DB 361 ACTTTAAGCGGGTGAAGTGAAATTTGGCTGTGATAAAGACATTTGAATCAACGCGAGTA 420
QY 579 AAGAAGTAACCGAACTTGAGGTTA CTAATGAAGGGCTTAAATCCTTGCATGTGGTAGCC 638
DB 421 AGAAGTTACCGAACTAGAAGTTACTAA CGAAGGACCTAAATCCTTGCATGTGGTAGCC 480
QY 639 ATTTCCACTTTCTTGAAGCTAACAGGCACTAAATTTGATCGTGAACCAAGCCATTTGGCA 698
DB 481 ATTTCCACTTTCTTGAAGCAACAAAGCAATTTGAATTCGATCGGGAACCAAGCCATTTGGCA 540
QY 699 AAGCCTAGATATTCCTCTGGCAACACGCTAGCAATTTGGGCGAGCAACCAAGCCGCAAG 758
DB 541 AAGCCTAGATATTCCTCTGGCAACACGCTAGCAATTTGGGCGAGCAACCAAGCCGTAAG 600
QY 759 TGCAGTTGATTCCTCTTGGTGGCAGTAAAGTGAATGTCATGAACGGCTTTGTGAATA 818
DB 601 TGCAGTTAATCCCTCTTGGCGGTAGTAAAGTGAATGTCATGAACGGCTTTGTGAATA 660
QY 819 ACATCGCGATGAACGCCATAACATAAAGCGCTTGCAAGGCGAAATCTCAGCGATTTA 878
DB 661 ATATTTGCGACGAACGCCATAACCAAGCGCTAGACAAAGCAAAATCTCAAGGATTTA 720
QY 879 TCAAGTAAAGGACTCCCATGAAATGAAACCAAGATATGTAATACCTACGACCC 938
DB 721 TCAAGTAAAGGACTCCCATGAAATGAAACCAAGATATGTAATACCTACGACCC 780
QY 939 ACCAAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 998
DB 781 ACCAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 840
QY 999 TATACCACTATGGCAAGAACCTTAAATTTGGCGCGGTAAACATACTATCCGTAGGGTATG 1058
DB 841 TATACCACTATGGCAAGAACCTTAAATTTGGGTGACAGGTAAACATACTATCCGTAGGGTATG 900
QY 1059 GGTGAGCAATAGCCCTGATGAACACCCCTAGATTTAGTCATCACTAACCGGATGATT 1118
DB 901 GGTGAGCAATAGCCCTGATGAACACCCCTAGATTTAGTCATCACTAACCGGATGATT 960
QY 1119 ATCGACTACACCGGATTTTAAAGCGGACATTTGGGATTTAAACCGGCAAAATCCATGGC 1178
DB 961 ATTGACTACCGGATTTTAAAGCGGACATTTGGGATTTAAACCGGCAAAATCCATGGC 1020
QY 1179 ATTGCAAGGCGAGGAACCAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1238
DB 1021 ATTGCAAGGCGAGGAACCAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1080
QY 1239 GTGGGCAAGAAGCACTACGAGGGAGGATGATTTATACCGCTGGGGGATCGATTCA 1298
DB 1081 GTGGGCAAGAAGCACTACGAGGGAGGATGATTTATACCGCTGGGGGATCGATTCA 1140
QY 1299 CACACCCACTTCTTTCTCCACAAATTTCCCTACCGCTCTAGCCAAATGGCGTTACCAAC 1358

DB 1141 CACACCCACTTCTCTCTCCACAACAATTCCTTACCCTCTAGCCAAATGGCGTTACAACA 1200
QY 1359 ATGTTTGGAGCGGCACAGAGTCTCTGTAGATGGCAGGAATGCGACTACTATCACTCCGGGC 1418
DB 1201 ATGTTTGGCGGTGGCACAGGCCCGTAGTAGTGCGACGAATGCGACTACCATCACTCGGGC 1260
QY 1419 AAATGGAACTTGCACCCGATGTTGCGGCGACGAGAAGATATTCTATGAATGTGGGCTTT 1478
DB 1261 AAATGGAACTTGCACCCGATGTTGCGGCGACGAGAAGATATTCTATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGCGCAATAGCTCTAGCAAAAACAACTTTGTAGAACAGTAGTAGAGCGGGCGG 1538
DB 1321 TTGGGCAAGCGCAATAGCTCTAGTAAAAACAACTTTGTAGAACAGTAGTAGAGCGGGCGG 1380
QY 1539 ATTGTTTAAATTTGCAATGAAGACTGGGGCAACCAACCAAGTGGCATCGATCACTGCTTTG 1598
DB 1381 ATTGTTTAAATTTGCAATGAAGACTGGGGCAACCTCCAGTGCATCGATCACTGCTTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCA 1658
DB 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACCCGATACGCTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCATGCCCTACCACTT 1718
DB 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCATGCCCTACCACTT 1560
QY 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGGAGGCGAGTCAATATT 1778
DB 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGGAGGCGAAGTGAATATT 1620
QY 1779 CTACCTCTCTCCACACCCCGCTATTTCCCTATACCAATTAATAGCGTTGCAGAACACTTTA 1838
DB 1621 CTACCTCTCTCCACACCCCGCTATTTCCCTATACCAATTAATAGCGTTGCAGAACACTTTA 1680
QY 1839 GACATGCTCATGACATGCCACCCCTAGACAAACCGCATCCGAGGAGTTTACAATTTCT 1898
DB 1681 GACATGCTTATGACCTGCCACCCCTAGATAAAGCATCCGAGGAGTCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTG 1958
DB 1741 CAAAGCGGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTG 1800
QY 1959 ATCGCATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGCAAGCGAAGTGAATCTCTCGA 2018
DB 1801 ATCGCATGACAAAGTTCGATTCGCAAGCAATGGGCGGTGCGGAGAGTGAATCTCTAGA 1860
QY 2019 ACTTGGCAGACTCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATGGCAAA 2078
DB 1861 ACTTGGCAACTCGCAGACAAAGAAATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATGGTGA 1920
QY 2079 GATAACGATAAATTCGCGATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTG 2138
DB 1921 GATAATGACAACTTCCGCTCAACCGCTATATCTCCAAATACACTATTAATCCCGCTTTG 1980
QY 2139 ACCACGGCTGAGGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2198
DB 1981 ACCATGGCTGAGGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2040
QY 2199 GTGTGGAATCCTGCTTTTTTGGCGTAAAAACCCAAAATTCGTGATCAAGGCGGTATGGTG 2258
DB 2041 GTGTGGAATCCTGCTTTTTTGGCGTAAAAACCCAAAATTCGTGATCAAGGCGGTATGGTG 2100
QY 2259 GTCTTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCACTCCCGCAACCGGTTATTATAC 2318
DB 2101 GTCTTCTGAAATGGCGGATTTCTAATTCGCTCTGTGCCCACTCTCTCAGCGGTTATTATAC 2160
QY 2319 CCGGAAATGTTTGGGATCACGCAAGCGGAAATTTTGACACAGCATCACATTTTGTGTCC 2378
DB 2161 CCGGAAATGTTTGGGATCACGCAAGCGGAAATTTTGACACAGCATCACATTTTGTGTCC 2220
QY 2379 AAAGTCCCTTATGAAATGGCGTGAAGAAAGCTGGGCTTTAGAGCGGCAAGTTCTTACCG 2438

Db	2221	AAAGTCGCCTATGAAATCGTGTGAAGAAAAAAGTGGTTAGAGCGCAAGGTGCTCCC	2280
Qy	2439	GTCAAAAACGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAA	2498
Db	2281	GTGAAAAACGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAA	2340
Qy	2499	ATCACCGTCGATCCGAAGACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA	2558
Db	2341	ATCACCGTCGATCCGAAGACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA	2400
Qy	2559	CCCACTCTCGAAGTGGCTCTAGCCCAACGCTACACTTTCTTAGGCACAAT	2610
Db	2401	CCCACTCTCGAAGTGGCTCTAGCCCAACGCTACACTTTCTTAGGCATAAT	2452
RESULT 5			
BD185303			
LOCUS	BD185303	2405 bp	DNA linear PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.		
ACCESSION	BD185303		
VERSION	BD185303.1	GI:31877503	
KEYWORDS	JP 2002355054-A/2.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	1 (bases 1 to 2405)		
TITLE	Kusters,J.G. and Cattoli,G.		
JOURNAL	Helicobacter felis vaccine		
COMMENT	Patent: JP 2002355054-A 2 10-DEC-2002; AKZO NOBEL NV		
	OS Helicobacter felis		
	PN JP 2002355054-A/2		
	PD 10-DEC-2002		
	PF 16-JUL-2001 JP 2001214711		
	PR 17-JUL-2000 EP 00202565.8		
	PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI		
	PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC A61K39/23,		
	PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,		
	PC C12N1/19,		
	PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC 53,		
	PC G01N33/53,G01N33/566,G01N33/569///(C12N9/80,C12R1:01),(C12Q1/68, PC C12R1:01),		
	PC C12N15/00,C12N5/00,A61K37/02		
	CC Helicobacter felis vaccine		
	PH Key Location/Qualifiers		
	FT CDS (1)..(681)		
	FT CDS (692)..(2398).		
FEATURES	Location/Qualifiers		
source	1..2405		
	/organism="Helicobacter felis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:214"		
ORIGIN			
Query Match	74.9%; Score 2158.6; DB 6; Length 2405;		
Best Local Similarity	93.6%; Pred. No. 0;		
Matches 2251; Conservative	0; Mismatches 154; Indels 0; Gaps 0;		
Qy	206	GTGAAACTCACCAAGAGCAAGAAAGTCTTGTATATATATGCGGGCGAAGTGGCT	265
Db	1	GTGAAACTCACCAAGAGCAAGAAAGTCTTGTATATATATGCGGGCGAAGTGGCT	60
Qy	266	AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT	325
Db	61	AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT	120
Qy	326	GCCCATATTATGAGCAAGCGCGGTGGAAAAAACCCTTGCCACGCTTATGGAAGAG	385

Db	121	GCCCATATTATGACAGGCGCGCGTGGCAAAAAAACCGTTGCTGAACTTATGGAAGAA	180
Qy	386	TGCATGCACATTTTGAATAAGATGAAGTAATCCCGGGGTGGTAAATATATGTTCCCGAT	445
Db	181	TGATATGCACATTTTGAATAAGATGAAGTAATCCCGGGGTGGTAAATATATGTTCCCGAT	240
Qy	446	CTAGGTGTAGAACCCACTTTCTGTATGTATGACGAACTTGTAACTGTGAATTTGGCCCATC	505
Db	241	TTGGGGGTAGAACCCACTTTCCCGATGGCAACCAACTCGTAACCGTGAATTTGGCCCATC	300
Qy	506	GAACCATGACGACACTTCAAAGCGGCGAAGTGAATTTTGGTTGCGATAAAGACATCGAG	565
Db	301	GAACCTGATGAACACTTTAAAGCGCGTGAAGTGAATTTTGGCTGTATTAAGACATTTGAG	360
Qy	566	CTCAATGCAGGCAAGAACTTAACCGAACTTTGAGGTTACTAATGAAGGCGCTTAAATCTCTTG	625
Db	361	CTCAACGCGGTAAAGAGTTTACCGAGCTTGAAGTTACCAACGAAGGACCTTAAATCTCTTG	420
Qy	626	CATGTGGGTAGCCATTTTCCACTTCTTTGAAGCTTAACAGGCACTTAAATTTGATGCTGAA	685
Db	421	CATGTGGGTAGCCATTTTCCACTTCTTTGAAACCAACCAAGGCAATTTGAAATTTGATGCTG	480
Qy	686	AAAGCCTATGCAAAACGCTAGATATTCCCTCTCGCAACGCTACGCAATTTGGGGCAGGA	745
Db	481	AAAGCCTATGCAAAACGCTAGATATTCCCTCTCGCAACGCTACGCAATTTGGGGCAGGA	540
Qy	746	CAAAACCGCAAAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC	805
Db	541	CAAAACCGCAAAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC	600
Qy	806	GGGCTTGTGAATAACATCGCGGATGAACCGCATTAACATAAAGCGCTTGACAGGCGGAAA	865
Db	601	GGGCTTGTGAATAATAATTTCGCGAGCAACGCGCATTAACATAAAGCGCTTGACAGGCGGAAA	660
Qy	866	TCTCACCGATTTATCAAGTAAAGGAGACTCCCATGAAGTAAAAAACAAGAAATATGTAAA	925
Db	661	TCTCACCGATTTATCAAGTAAAGGAGACTCCCATGAAGTAAAAAACAAGAAATATGTAAA	720
Qy	926	TACCTAGCGGACCCACCAAGCGATTAAGTGGCTTAGGAGATACCGATCTTTTGGGCGAGA	985
Db	721	CACCTACGGACCCACCAAGCGATTAAGTGGCTTAGGAGATACCGATCTTTTGGGCGAGA	780
Qy	986	AGTAGAATCATGATATATACCATATGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT	1045
Db	781	AGTAGAATCATGATATATACCATATGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT	840
Qy	1046	CCGTAGGGTATGGTTCAGAGCAATAGCCCTGATGAAAACACCCCTAGATTTTATGTCATCAC	1105
Db	841	CCGTAGGGTATGGTTCAGAGCAATAGCCCTGATGAAAACACCCCTAGATTTTATGTCATCAC	900
Qy	1106	TAAACCGATGATTTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGG	1165
Db	901	CAACCGATGATTTATCGACTACACCGGATTTTATTAAGCCGACATTTGGTATTTAAAAATGG	960
Qy	1166	CAAAATCCATGGCATTTGCAAGGCAAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA	1225
Db	961	CAAAATCCATGGCATTTGCAAGGCGGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA	1020
Qy	1226	TATGTCGTGGGTGTGGGCACAGAGCACTAGCAGGGGAAGTATGATTTATACCGCTCG	1285
Db	1021	TATGTCGTGGGTGTGGGCACAGAGCACTAGCAGGGGAAGTATGATTTATACCGCTCG	1080
Qy	1286	GGGAATCCATTTACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCCAA	1345
Db	1081	GGGAATCCATTTACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCCAA	1140
Qy	1346	TGGCGTTACAAACCATGTTTGGAGGCGGCACAGGTCCTGTAGATGGCAGCAATGCGACTAC	1405
Db	1141	TGGTGTACAAACCATGTTTGGAGGTCGCACAGGTCGGGTAGATGGCAGCAATGCGACCAC	1200
Qy	1406	TATCATCTCCGGGCAATGGAACCTTGACCGGATGTTGGCGCAGCAGAGATTTCTAT	1465

Db 1201 CATCACTCGGCGCAATGGAACCTTGACCGCATGTTGCGCGAGCTGAAGAGTATTCTAT 1260
Qy 1466 GAATGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAACAACACTCGTAGAACAGT 1320
Qy 1526 AGAAGGGGCGGATGTTGGTTTAAATTGATGAAGACTGGGGCAACAACCAAGTGGAT 1585
Db 1321 AGAAGGGGCGGATGTTGGTTTAAATTGATGAAGACTGGGGCAACAACCAAGTGGAT 1380
Qy 1586 CGATCACTCTTGAGCGTGCGAGATGAATACCATGTGCAAGTTTGTATCCACCGATAC 1645
Db 1381 CGATCACTCTTGAGCGTGAGAGATGAATACCATGTGCAAGTTTGTATCCACCGATAC 1440
Qy 1646 AGTCAATGAGGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGGCCATCA 1705
Db 1441 GGTCAATGAGGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGGCCATCA 1500
Qy 1706 TGCCTACCACTGAGGAGGCGGTGGAGGACATCACTCATGTATGTTATCAATGGCAGG 1765
Db 1501 TGCCTACCACTGAGGAGGCGGTGGAGGACATCACTCATGTATGTTATCAATGGCAGG 1560
Qy 1766 CGAGCTCAATATTCTACCTCTCTCCACACCCCACTATTTCCTATACCATTAATACGCT 1825
Db 1561 CGAGCTCAATATTCTACCTCTCTCCACACCCCACTATTTCCTATACCATTAATACGCT 1620
Qy 1826 TGCAGAACCTTAGACATGCTCATGACATGCGACACCTAGACAAACGATCCGCGAGA 1885
Db 1621 TGCAGAACCTTAGACATGCTCATGACATGCGACACCTAGACAAACGATCCGCGAGA 1680
Qy 1886 TTTACAATTTTCTCAAAGCGTATCGGCCCGGCTCTATCGGGCTTGAAGATGTGTCCA 1945
Db 1681 TCTCCAGTTTCTCAAAGCGTATCGGCCCGGCTCTATCGGGCTTGAAGATGTGTCCA 1740
Qy 1946 TGATATGGGTGTGATCGGATGACAGCTCGGATTCGGAAGCATGGGGCGGTGCGAGCGA 2005
Db 1741 TGATATGGGTGTGATCGGATGACAGCTCGGATTCGGAAGCATGGGGCGGTGCGAGCGA 1800
Qy 2006 AGTGATTCCTCAACTTGTGCACACTGCGGATGAAGATAAAGAAATTTGGTAACTTCC 2065
Db 1801 AGTGATTCCTCAACTTGTGCACACTGCGGATGAAGATAAAGAAATTTGGTAACTTCC 1860
Qy 2066 TGAAGATGCCAAGATACGATAAATTTCCGCAATTAAAGCGCTACATCTCCAAATACACTAT 2125
Db 1861 TGAAGATGGTCAGATATGACAACTTCGCGATCAACCGCTATATCTCCAAATACACTAT 1920
Qy 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGATATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1921 TAATCCCGCTTTGACCCAGCGGTGAGCGATATATCGGCTCTGTGGAAGAGGGCAAGAT 1980
Qy 2186 CGCGACTTGTGGTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGACTTGTGGTGTGGAATCCTGCTTTTGGTGTAAACCCCAAAATCGTGATCAA 2040
Qy 2246 AGCGGTATGGTGGTCTTCTCTGAATGGCGGATTTCTAACCGGTCTGTGCCACATCCCA 2305
Db 2041 AGCGGTATGGTGGTCTTCTCTGAATGGCGGATTTCTAACCGGTCTGTGCCACATCCCA 2100
Qy 2306 ACCGGTTTATACCGGCAAAATGTTTGGGCATCACGGCAAGGGCGAAATTTGACACAGCAT 2365
Db 2101 GCGGTTTATACCGGCAAAATGTTTGGGCATCACGGCAAGGGCGAAATTTGACACAGCAT 2160
Qy 2366 CACTTTGTTTCCAAAGTGCCTATGAAAATGGCGTGAAGAAAGCTGGGCTTAGAGG 2425
Db 2161 CACTTTGTTTCCAAAGTGCCTATGAAAATGGCGTGAAGAAAGAACTAGGCTTAGAGG 2220
Qy 2426 CCAAGTCTACCGGTCAAAAATGCGCGTAAACATCAACCAAGAAAGCTTCAAGTCAACGA 2485
Db 2221 CAAGTGTCTACCGGTGA AAAATGCGCGCAACATCACTAGAAGAAAGCTTCAAAATCAACAA 2280
Qy 2486 CAAAACGGCAAAAATCACCGTGCATCCGAAAAACCTTTCGAGGTCTTTGTAGATGGCAACT 2545
Db 2281 CAAGACGGGCATATCACTGTGATCCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT 2340

Qy 2546 CTGCACCTCTAAACCCCACTCGCAAGTGCGCTCTAGCCACCGCTACACTTCTTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCGCTCTAGCCACCGCTACACTTCTTCTAGGC 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405
RESULT 6
LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 4 from Patent EP1176192.
ACCESSION AX356686
VERSION AX356686.1 GI:18674023
KEYWORDS Helicobacter felis
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 4 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
source
1..2405
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
CDS
1..681
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23257.1"
/db_xref="GI:18674024"
/translation="MKLTPEQEKFLLYVAGEVARKKAEGLKLNQPEAIAYISAHIM
DEARRGKTVAELMBECMHFLKDEVMGVGMVVDLGVATFPDGTGLVTVMVPIEP
DEHFGAVEYKFCDDIELNAGKEVTELVNKGPKSLHVGSHFPEETNKALKFDRK
KAYGRKLDIPSGNTLIGAGTRKQVLIPLGGSKKVIQMGVLVNNIADERHKHKKALDK
AKSHGFIK"
692..2398
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23258.1"
/db_xref="GI:18674025"
/translation="MKMKQEVVNTYGPTTGDVKRLGDTDLMAEVEHDYTTYGBELKF
GAGKTIREGMGQSNPDENTDLVITNMIIDYGIYKADIGIKNGKIHGIVKAGND
MODGVSPIHVVGVGTALAGEGMIITAGGIDSHTEFLSPQFPPTALANGVTTMFGGT
GPDVGNATTITPGKNLHRLRAAEEYSNMVGFLOGKSSSKQLVEQVEAGAIGPK
LHEDWGTPSAIDHCLSVADYDVQVCIHTDVNEAGYVDDTLNMGRAIHAHYHSG
ACGHSIPDVTWAGELNIPSSSTPTPTINTVAEHLDMLTCHLDKRIREDLPQK
QSRIRPGSIAEDVLHDIAMTSSDSQAMGAVEIPRTWQADKNKKEFGKLPED
GADNFRITKYISKTINPALTHGVSEYIGSVGEKIALDVVMNPAFVGVPKIVIK
GEMVFSMGSDSNASVPTPQPVYREMFGHGKAKFDTSLTFVSKVAYENGVEKELGI
ERKLPVKNCRNITTKDFKFNKNTAHITVDPFTFEVDGKLCITSPASEVPLAQRIT
F"

ORIGIN

Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred.No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
Qy 206 GTGAAACTCACACCCCAAGAGCAAGAAAAAGTCTTGTATTATTATGCGGGCGAAGTGCGCT 265
Db 1 GTGAAACTCACACCCCAAGAGCAAGAAAAAGTCTTGTATTATTATGCGGGCGAAGTGCGCT 60
Qy 266 AGAAGCGCAAGAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGCGCAAGAGCAGAGGGCTTAAAGCTCAATCAACCCCAAGCCATTGCTTACATTAGT 120

QY 326 GCCATATTTATGACGAAGCGCCGTTGNAAAAAAACCGTTCCCGAGCTTATGGAAGAG 385
Db 121 GCCCATATTTATGACGAGCGCCGTTGGCAAAAAAACCGTTGCTGAACCTTATGGAAGAA 180
QY 386 TGCATGCATTTTGTAAAAAAGATGAAGTAATCCCGGGGTGGTAAATATGTTTCCCGAT 445
Db 181 TGTATGCATTTTGTAAAAAAGATGAGTGATGCCCGGTGTGGGAATATGTTCCCTGAT 240
QY 446 CTAGTGTATGAAGCCACTTTCTCTGATGTGTCGAAACTTTGTAACTGTGAATTTGGCCCATC 505
Db 241 TTGGCGGTAGAAGCCACTTTCCCGGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAAT 300
QY 506 GAAACAGATGAGCACTTCAAGCGGGGGAAGTGAATTTGGTTGGCATTAAGAATCATCGAG 565
Db 301 GAACTGTATGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGTATAAAGACATTGAG 360
QY 566 CTCAATGCGAGCAAGTAACCGAACTTTGAGGTTACTAATCAAGGGCCCTAAATCCTTG 625
Db 361 CTCAACGCGGTAAAGGATTTACCGAGCTTTGAAGTTTACCACGAAGAACCTAAATCCTTG 420
QY 626 CATGTGGGTAGCCATTTTCCACTTTCTTTGAAGCTAAACAAGGCACCTAAAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTTCCACTTTCTTTGAACCAACAAGGCAATTTGAAATTCGATCGGAA 480
QY 686 AAAGCCTATGGCAAAACGCTTAGATATATCCCTCTGGGCAACAGCTACGATTTGGGGCAGGA 745
Db 481 AAAGCCTATGGCAAAACGCTTAGATATATCCCTCTGGCAACAGCTACGATTTGGGGCAGGA 540
QY 746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAGATGATTTGCCATGCAAC 805
Db 541 CAACCCGCAAGTGCAGTTGATTCCTCTTGGCGGTAGTAAAAAAGTGAATTTGCCATGAAC 600
QY 806 GGGCTGTGTAATAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGAACAAGGGGAAA 865
Db 601 GGGCTGTGTAATAATTTGGGACGACGCGCATAAACACAAGCACTAGACAGGCAAAA 660
QY 866 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAA 925
Db 661 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAGTATGTAAA 720
QY 926 TACCTACGGACCCACCAAGCGGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 985
Db 721 CACCTACGGACCCACCAAGCGGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 780
QY 986 AGTAGAACATGACTATACCACTATATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 781 AGTAGAACATGACTATACCACTATATGGCGAAGACTTAAATTTGGCGCGGTAAAACTAT 840
QY 1046 CCGTAGGGGTATGGGTGAGAGCAATAGCCCTGTAGAAAAACCCCTAGATTTAGTCAATC 1105
Db 841 CCGTAGGGGTATGGGTGAGAGCAATAGCCCAAGATGAAAAACACCTTAGATTTAGTCAATC 900
QY 1106 TAACCGGATGATTTATGCACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAACGG 1165
Db 901 CAACGCGATGATTTATGCACTACACCGGATTTATAAAGCCGACATTTGGTATTTAAAAATGG 960
QY 1166 CAAAAATCCATGGCATTTGCAAGGAGGAGGAAACAAGGACATGCAAGTGGCTTAAGCCCTCA 1225
Db 961 CAAAAATCCATGGCATTTGCAAGGAGGAGGAGGAAACAAGATGCAAGTGGCTTAAGCCCTCA 1020
QY 1226 TATGGTCTGGGTGTGGGCACAGAGCACTAGCAGGGGAAGGTAATGATTTATACCGCTGG 1285
Db 1021 TATGGTCTGGGTGTGGGCACAGAGCACTAGCAGGGGAAGGTAATGATTTATACCGCTGG 1080
QY 1286 GGGATCGATTTACACACCACTTCTCTTCTCAACAATTTCCCTACCGCTCTAGGCCAA 1345
Db 1081 GGGATCGATTTGACACACCACTTCTCTCTCTCCCAACAATTTCCCTACCGCTCTAGGCCAA 1140
QY 1346 TGGGTTTACACCAATGTTTGGAGCGGCACAGGTCCTGTAGATGCGCAATGCGACTAC 1405
Db 1141 TGGGTTTACACCAATGTTTGGAGTGGCACAGGTCCTGTAGATGCGCAATGCGACTAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAACTTTGCACCGCATTTGTGGCGCAGCAGAGATTTCTAT 1465

Db 1201 CATCACTCCGGGCAAAATGGAACCTTGACCCGATTTGTCGCGCAGCTGAAAGATTTCTAT 1260
QY 1466 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCCAGTAAAAAACAATCTGTAGAACAGT 1320
QY 1526 AGAAGCGGGCGCATTTGGTTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGGAT 1585
Db 1321 AGAAGCGGGCGCATTTGGTTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGGAT 1380
QY 1586 CGATCACTGCTTGCAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGGATAC 1645
Db 1381 CGATCACTGCTTGCAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGGATAC 1440
QY 1646 AGTCAATAGCGCAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCATCCA 1705
Db 1441 GGTCAATAGCGCAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCATCCA 1500
QY 1706 TGCTTACCAATTTGAGGGAGCGGTGGAGACACTCACTGATGTTATCAACATGGCAGG 1765
Db 1501 TGCTTACCAATTTGAGGGAGCGGTGGAGACACTCACTGATGTTATCAACATGGCAGG 1560
QY 1766 CGAGCTCAATTTCTACCCCTCTCCACCCACCTTATCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGCTCAATTTCTACCCCTCTCCACCCACCTTATCCCTATACCAATTAATACGGT 1620
QY 1826 TGCAAGAACACTTAGACATGCTCATGATGACACCACTAGACAAAAGCATCCGCGAGGA 1885
Db 1621 TGCAAGAACACTTAGACATGCTCATGATGACACCACTAGACAAAAGCATCCGCGAGGA 1680
QY 1886 TTTTCAATTTTCTCAAAAGCGGTATCCCGCCGGCTCTATCCGGCTGAAGATGTGCTCA 1945
Db 1681 TCTCCAGTTTTTCCAAAGCGGTATCCCGCCGGCTCTATTTCCCGCTGAAGATGTGCTCA 1740
QY 1946 TGATATGGGTGTGATCCGATGACAGCTCGGATTCGCAAGCAATGGCGCTGAGGCGA 2005
Db 1741 TGATATTTGGGTGTGATCCGATGACAGCTCGGATTCGCAAGCAATGGCGCTGAGGCGA 1800
QY 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGATAAAAAAGAAATTTGTAAGCTTCC 2065
Db 1801 AGTGATTTCTTAGAACTTTGGCAACTGCAGACAGAAATAAAAAGAAATTTGTAAGCTTCC 1860
QY 2066 TGAAGATGGCAAAAGATAACGATAATTTCCGATTTAAGCGCTACATCTCAAAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATAATGACAACTTCCGATCAAAAGCTATATCTCCTCAAAATACACTAT 1920
QY 2126 CAACCCCGCTTTGACCCACCGCGTGGAGGATATCGGCTCTGTGGAGAGGCGCAAGAT 2185
Db 1921 TAATCCCGCTTTGACCCATCGCGTGGAGGATATCGGCTCTGTGGAGAGGCGCAAGAT 1980
QY 2186 CGCCGACTTGGTGTGGAAATCCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1981 CGCCGACTTGGTGTGGAAATCCTGCTTTTGGTGTAAAAACCAAAATCGTGATCAA 2040
QY 2246 AGCGGTATGGTGTCTTCTCTGAAATGGGCGATTTTAAACCGCTGTGTCACCTCCCA 2305
Db 2041 AGCGGTATGGTGTGTTCTCTGAAATGGGCGATTTAAACCGCTGTGTCACCACTCA 2100
QY 2306 ACCGGTTTATTAACCGGAAATGTTTGGGATCAAGGAGGCGAAATTTGACACCAAGAT 2365
Db 2101 GCGCGTTTATTAACCGGAAATGTTTGGGATCAAGGAGGCGAAATTTGACACCAAGAT 2160
QY 2366 CACTTTTGTTCCAAAGTCGCTTATGAAATGGGCTGAAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCAAAGTCGCTTATGAAATGGGCTGAAAGAAAGCTAGGCTTAGAGCG 2220
QY 2426 CCAAGTTCTACCGGTCAAAAAAGCTCCGCTTAACATCAACAAAGAAAGCTTTCAAGTTCAACGA 2485
Db 2221 CAAAGTGTACCGGTGAAAAAGCTCCGCAACATCACTAAGAAAGACTTTCAAATTCACAA 2280
QY 2486 CAAACCGGCAAAATCAACCGTCAATCCGAAACCTTTCAGGCTTTTGTAGATGGCAACT 2545

Db	2281	CAAGACGGCGCATATCACCTGTGCGATCCTTAAACCTTCGAGGTCCTTTGTAGATGGCAACT	2340
Qy	2546	CTGCACCTCTAAACCCACCTCGCAAGTGCTCTAGCCAGCGCTACACTTTCTTCTAGGC	2605
Db	2341	CTGCACCTCTAAACCGGCTCTGAAGTGCTCTAGCCCAACGCTACACTTTCTTAGGC	2400
Qy	2606	ACAAAT 2610	
Db	2401	ACAAAT 2405	
RESULT 7			
BD185305			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
OS			
PN			
PD			
PF			
PR			
PI			
PC			
C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC			
A61K39/23,			
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,			
PC C12N1/19,			
PC			
C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC			
53,			
PC			
G01N33/53,G01N33/566,G01N33/569// (C12N9/80,C12R1:01),(C12Q1/68, PC			
C12R1:01),			
PC C12N15/00,C12N5/00,A61K37/02			
CC Helicobacter felis vaccine			
FH Key Location/Qualifiers			
FT CDS (2)..(682)			
FT CDS (693)..(2399).			
FEATURES			
source			
1..2407			
/organism="Helicobacter felis"			
/mol_type="genomic DNA"			
/db_xref="taxon:214"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2251; Conservative			
0; Mismatches 155; Indels			
0; Gaps			
0;			
Qy	206	GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTTATATTATGCGGGCGAAGTGCT	265
Db	2	GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTTATATTATGCGGGCGAAGTGCT	61
Qy	266	AGAAAGCGAAGCGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT	325
Db	62	AGAAAGCGAAGCGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT	121
Qy	326	GCCCATATTATGACGAAGCGCGCTGGAAGAAACCCGTTGCCAGCTTATGGAAGAG	385
Db	122	GCCCATATTATGACGAAGCGCGCTGGAAGAAACCCGTTGCCAGCTTATGGAAGAG	181
Qy	386	TGCATGCATTTTGAAGAAAGATGAAGTAATATGCCCGGGGTGGGTAAATATGTTTCCCGAT	445

Db	1262	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAATAAACAACATCTGTAGAACAAAT	1321
Qy	1526	AGAAGCGGGCGGCGATTTGTTTAAATTTGCAATGAAGACTGGGGCACAAACACCAAGTGCAT	1585
Db	1322	AGAAGCGGGCGGATCGGCTTTAAATTTGCATGAAGACTGGGGCACAACTCCAAGTGAAT	1381
Qy	1586	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGATTCACACCGATAC	1645
Db	1382	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGATTCACACCGATAC	1441
Qy	1646	AGTCAATAGGCGAGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCCATCCA	1705
Db	1442	GGTCAATAGGCGAGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCCATCCA	1501
Qy	1706	TGCTTACCACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG	1765
Db	1502	TGCTTACCACATTGAGGAGCGGGGAGGACACTCACCTGATGTTTATCACCATGGCAGG	1561
Qy	1766	CGAGCTCAATATTCTACCTCTCCACCAACCCGACATATTCCTTATACCATTAATACGGT	1825
Db	1562	CGAGCTCAATATTCTACCTCTCCACCAACCCGACATATTCCTTATACCATTAATACGGT	1621
Qy	1826	TGCAGAACCTTAGACATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA	1885
Db	1622	TGCAGAACCTTAGACATGCTCATGACCTGCCACCTAGATAAAGCATCCGCGAGGA	1681
Qy	1886	TTTACAAATTTCTCAAAGCGGTATCCGCCCGGCTCTATCCCGGCTGAAGATGTCTCCA	1945
Db	1682	TTTACAAATTTCTCAAAGCGGTATCCGCCCGGCTCTATCCCGGCTGAAGATGTCTCCA	1741
Qy	1946	TGATATGGGTGTGATCCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGGTGAGCGGA	2005
Db	1742	TGATATGGGTGTGATCCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGGTGAGCGGA	1801
Qy	2006	AGTGATTCCTCGAATCGGAGCTCGGATGAGATAAAGAAATTTGGTAAGCTTCC	2065
Db	1802	AGTGATTCCTCGAATCGGAGCTCGGATGAGATAAAGAAATTTGGTAAGCTTCC	1861
Qy	2066	TGAAGATGGCAAGATGAACATATTTCCGATTAAGCGCTACATCTCCAAATACACTAT	2125
Db	1862	TGAAGATGGCGAGATGAACACACTTTCGATTCACAAAGCTACATCTCCAAATACACTAT	1921
Qy	2126	CAACCCCGCTTTGACCCACCGCGTAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGT	2185
Db	1922	TAACCCCGCTCTAACCCATCGGGTAAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT	1981
Qy	2186	CGCGACTTGGTGTGGTAACTCTGCGCTTTTGGCGTAAACCCCAAAATCGTGATCAA	2245
Db	1982	CGCTGATTTGGTGTGGTAACTCTGCGCTTTTGGCGTAAACCCCAAAATCGTGATCAA	2041
Qy	2246	AGCGGTATGGTGTCTCTGAAATGGGCGATTTCTAACGCGTCTGTGCCACTCCCCCA	2305
Db	2042	AGCGGTATGGTGTCTCTGAAATGGGCGACTCCAACGCGTCCGTGCGCTACACCTCA	2101
Qy	2306	ACCGGTTTATTACCGGAAATGTTTGGGCATCACGCAAGCGCGAAATTTGACACCGCAT	2365
Db	2102	GCCGGTTTATTACCGGAAATGTTTGGGCATCACGCAAGCGCGAAATTTGACACCGCAT	2161
Qy	2366	CACCTTTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCG	2425
Db	2162	CACCTTTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTTAGAGCG	2221
Qy	2426	CCAAAGTTCTACCGTCAAAACCTCCGCTAACATCACCAAGAAAGACTTCAAGTTCAACGA	2485
Db	2222	CAAGGTGCTACCGTCAAAACCTCCGCGAACATCACTAAGAAAGACTTCAAAATTCACAA	2281
Qy	2486	CAAAACGGCAAAATACACCGTCTGATCCGAAACCTTTCGAGTCTTTGTGATGCGCAAACT	2545
Db	2282	CAAGACGGCGCATATCACTGTGATCTCTAAACCTTCGAGTCTTTGTGATGCGCAAACT	2341
Qy	2546	CTGACCTCTAAACCCACCTCGAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGGC	2605
Db	2342	CTGACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGGC	2401

Qy	2606	ACAATG 2611	
Db	2402	NCAATG 2407	
RESULT 8			
AX356692		2407 bp	DNA
LOCUS		Sequence 10 from Patent EP1176192.	linear PAT 13-FEB-2002
DEFINITION			
AX356692			
ACCESSION			
AX356692.1		GI:18674029	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Helicobacter felis			
Helicobacter felis			
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;			
Helicobacteraceae; Helicobacter.			
REFERENCE			
1			
Kusters, J.G. and Cattoli, G.			
Helicobacter felis vaccine			
Patent: Ep 1176192-A 10 30-JAN-2002;			
Akzo Nobel N.V. (NL)			
FEATURES			
source			
1..2407			
/organism="Helicobacter felis"			
/mol_type="unassigned DNA"			
/db_xref="taxon:214"			
2..682			
/note="unnamed protein product"			
/codon_start=1			
/transl_table=11			
/protein_id="CAD23261.1"			
/db_xref="GI:18674030"			
/translation="MKLTPKEQKFLLYAGAEVARKKAEGLKLNQPEAIAYISAHIM			
DEARRGKKTVAELMEECMHFLKDDVMPGVGNMVPDLGVEATFDGTKLVTVNVP			
DEHFKAGEVPGCKDIELNAGKEVTELENTNEGPKSLHVGSHHFFPANKLKFDE			
KAYGRUDIPSGNTRLRIGAGTRKVKQLIPGSSKKVIGMNLVNNIADIRHKKALEK			
AKSHGFIK"			
693..2399			
/note="unnamed protein product"			
/codon_start=1			
/transl_table=11			
/protein_id="CAD23262.1"			
/db_xref="GI:18674031"			
/translation="MKMKQEVNVTYGTGDKVRLGDTDLWAEVEHDDYTTVEELKF			
GAGKTIREGMQSNSPDENTLDLVITNAMIIDYTGIVKADIGIKNGKIHGIGKAGNKD			
MODGVSPIVMVVGTEALAGEMLITAGGIDSHFLSPQOPPTALANGVTTMFGCGT			
GPVDGNATTITPGKWNLRMLRAAEYSNNVFLGKGNSSKKQLVQIEAGAIQPK			
LHBDWGTTPSAIDHCLSVADVDQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEG			
AGGHSPIVTIMAGELNLPSSSTPTTPTTINTVAEHLMLTCHLDRIRREDIQFS			
QSRIRPGSIAAEDVLHDIGVIAMTSSDQAMGRAGEVTPRTQWADKNKGFGLPED			
SANDNFRIKRYISKYITNPALTHGVSEYIGSVBEKIALDLVWNPAPFGVKPKIVIK			
GMVVPFEMGSDNASVTPQPVVYREMGHHGKAKFDTSITFVSKVAVENGVEKLG			
ERKVLVPVKNCRNITKDFKFNNTAHITVDPKTFEVDGKLCSTSKPASEVPLAQR			
YF"			
ORIGIN			
Query Match	74.9%;	Score 2158.6;	DB 6; Length 2407;
Best Local Similarity	93.6%;	Pred. No. 0;	
Matches 2251;	Conservative 0;	Mismatches 155;	Indels 0; Gaps 0;
Qy	206	GTGAAACTCACACCAAGAGCAAGAAAGTCTTGTGTATATATATCGCGGCGAAGTGGCT	265
Db	2	GTGAACTCACACCAAGAGCAAGAAAGTCTTGTGTATATATATCGCGGCGAAGTGGCT	61
Qy	266	AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTACATTAGT	325
Db	62	AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTACATTAGT	121
Qy	326	GCCCATATTAGGACGAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGAAGAG	385
Db	122	GCCCATATTAGGACGAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGAAGAG	181

QY	386	TGATGACACTTTTGGAAAAAGATGAAGTAATGCCCCGGGGTGGGTAATATATGTTTCCCGAT	445
DB	182	TGTATGACACTTTTGGAAAAAGACGAGGTGATGCCCCGGTGTGGGTAATATGTTCCCTGAT	241
QY	446	CTAGGTGTAGAGCCACCTTCTCTGATGTGACGAACCTTGTAACTGTGAATTTGGCCCATC	505
DB	242	TTAGGCGTGAAGCTACTTTTCCGATGCGACCAAACTCGTAACCGGTGAATTTGGCCCATC	301
QY	506	GAACCAAGATGAGCACTTTCAAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG	565
DB	302	GAACCCGATGAACACTTTCAAAGCGGCGAAGTCAAAATTTGGCTGTGATTAAGACATTGAA	361
QY	566	CTCAATGCGAGGCAAGAAGTAACCGAACTTTGAGGTTTACTAAAGAGGCGCTAAATCCTTG	625
DB	362	CTCAACCGCAGGTAAAGCAAGTTACCGAACTTAGAAGTTTACCAAAGAGCACTAAATCCTTG	421
QY	626	CATGTGGGTAGCACTTTCCACTTCTTTTGAAGCTTACAAAGGCATTAATTCGATCGTGA	685
DB	422	CATGTGGGTAGCACTTTCCACTTCTTTTGAAGCCAAACAAAGGCATTTGAATTCGATCGGAA	481
QY	686	AAAGCCTATGGCAAAAGCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA	745
DB	482	AAAGCCTATGGCAAGCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA	541
QY	746	CAAAACCGCAAAAGTGCAGTTGATTCCTTTGGTGGCAGTAAAAAAGTGAATGGCATGAAC	805
DB	542	CAAAACCGTAAAGTGCAGTTAATTCCTCTTTGGCGCAGTAAAAAAGTGAATGGCATGAAC	601
QY	806	GGGCTTGTGAATAACATCCGGATGAACGCCATAAACAATAAAGCGTTGACAAGGGCAAA	865
DB	602	GGGCTTGTGAATAATTCAGATGAACGCCATAAACAACAAGCGTTAGAAAAAGCAAAA	661
QY	866	TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGATAATGTAAA	925
DB	662	TCTCAGCGATTTATCAAAATGAAGAGACTCCCATGAAAAATGAAAAACAAGATATGTAAA	721
QY	926	TACCTACGGACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA	985
DB	722	TACCTACGGACCTTACCACAGGCGACAAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA	781
QY	986	AGTAGAACATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT	1045
DB	782	AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT	841
QY	1046	CGGTGAGGATGGGTACAGCAATAGCCCTCATGAACAACCCATAGATTTAGTCATCAC	1105
DB	842	CGGTGAGGCGATGGGTACAGCAATAGTCCAGATGAACCAACCTAGATTTAGTCATCAC	901
QY	1106	TAAACGCGATGATATTCGACTACACCGGGAATTTACAAAGCCGACATTTGGGATTAATAACGG	1165
DB	902	CAACGCGATGATTTAGTACTACACCGGATTTTACAAAGCCGACATTTGGCAATTAATAATGG	961
QY	1166	CAAAATCCATGGCATTTGGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCA	1225
DB	962	CAAAATCCATGGCATTTGGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCA	1021
QY	1226	TATGTCTGGGTGTGGGACAGAGACTACACAGGGAAGTATGATTTATACCGCTGG	1285
DB	1022	TATGTCTGGGTGTGGGACAGAGACTATAGCAGGGGAAGTATGATTTATACCGCTGG	1081
QY	1286	GGGAATCGATTCACACACCCACTTCTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA	1345
DB	1082	GGGATCGATTCACACACCCACTTCTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA	1141
QY	1346	TGGCGTTACAACCATGTTTGGAGGGGCGCACAGGTCCCTGTAGATGGCACGAATGCGACTAC	1405
DB	1142	TGGCGTTACAACCATGTTTGGCGGTGGCACAGGTCCGCTAGATGGCACGAATGCGACTAC	1201
QY	1406	TATCACTCCGGCAATGGAATTCGACCGCATGTTGGCGGACGACGAAGATATTTCTAT	1465
DB	1202	CATCACTCCGGCAATGGAATTCGACCGCATGTTGGCGGACGATGGAAGATATTTCTAT	1261
QY	1466	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAACT	1525

DB	1262	GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCCAGTAAAAAAACAACCTTTGTAGAACAAAT	1321
QY	1526	AGAAGCGGGCGGCAATTTGAAATTTGCAATGAAGACTGGGGCACAAACCAAGTGGCAAT	1585
DB	1322	AGAAGCGGGCGGCAATTTGAAATTTGCAATGAAGACTGGGGCACAACTCCNAGTGAAT	1381
QY	1586	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATAC	1645
DB	1382	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATAC	1441
QY	1646	AGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCA	1705
DB	1442	GGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCA	1501
QY	1706	TGCTTACCACATTTGAGGAGCGGTTGGAGGACACTCACTCATGTGTTTATCACCATGGCAGG	1765
DB	1502	TGCTTACCACATTTGAGGAGCGGTTGGAGGACACTCACTCATGTGTTTATCACCATGGCAGG	1561
QY	1766	CGAGCTCAATATTTCTACCCCTCTCTCCACCCACCCCTATTTCCCTATACCAATTAATACGGT	1825
DB	1562	CGAGCTCAATATTTCTACCCCTCTCTCCACACCCCTATTTCCCTATACCAATTAATACGGT	1621
QY	1826	TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAAGCGATTCGCGAGGA	1885
DB	1622	TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGATAAAACGCGATTCGCGAGGA	1681
QY	1886	TTTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGGAAGATGTGCTCCA	1945
DB	1682	TTTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGGAAGATGTGCTCCA	1741
QY	1946	TGATATGGGTTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGA	2005
DB	1742	TGATATTTGGGTTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGA	1801
QY	2006	AGTGAATTCCTCGAACTTTGGCAGACTCGGATTAAGATAAAAAAGAAATTTGGTAAGCTTCC	2065
DB	1802	AGTGAATTCCTCGAACTTTGGCAGACTCGGATTAAGATAAAAAAGAAATTTGGTAAGCTTCC	1861
QY	2066	TGAAGATGCAAGAATAACGATTAATTTCCGATTAAGCGCTATACCTCCCAANTACACTAT	2125
DB	1862	TGAAGATGCTAGATTAACGATTAATTTCCGATTAACAAACGCTATCTCCAAATACACTAT	1921
QY	2126	CAACCCGCTTTGACCCACCGGTGAGCGATATATCGGCTCTGTGGAAGAGGCGCAAGAT	2185
DB	1922	TAACCCGCTCTTAAACCCATGGGTAGCGATATATCGGCTCTGTGGAAGAGGCGCAAAAT	1981
QY	2186	CGCCGACTTTGGTGGTGGAAATCCTGCTTTTGGCGTAAAAACCCAAAAATCGTGATCAA	2245
DB	1982	CGCTGATTTGGTGGTGGAAATCCTGCTTTTGGTGTGAAACCTTAAGATTTGTGATCAA	2041
QY	2246	AGGCGTATGGTGGTCTTCTCTGAAATGGGCGATTTCTAACGGGTCTGTGCGCACTCCCCA	2305
DB	2042	AGGCGTATGGTGGTCTTCTCTGAAATGGGCGCATCCCAACGGCTCGCTACACCTCA	2101
QY	2306	ACGGTTTATTAACCGGAAATGTTTGGGCGATCAACGCAAGGCGAAATTTGACACCCAGCAT	2365
DB	2102	GCGGTTTATTAACCGGAAATGTTTGGGCGATCAACGCAAGGCGAAATTTGACACCCAGCAT	2161
QY	2366	CACCTTTGTTTCCAAAGTCGCTTATGAAATGGCGTGAAGAAAAAGCTTGGGCTTTAGAGCG	2425
DB	2162	CACCTTTGTTTCCAAAGTCGCTTATGAAATGGCGTGAAGAAAAAGCTTGGGCTTTAGAGCG	2221
QY	2426	CCAAGTTCTACCGGTCAAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA	2485
DB	2222	CAAGGTGCTACCGGTGAAAAAATGCGCGCAACATCACTAAGAAAAAGACTTCAAAATTCACAA	2281
QY	2486	CAAAACGCAAAAAATCACCGTGCATCCGAAACCTTTCGAGGTCTTTGTAGATGCGAACT	2545
DB	2282	CAAGACGCGGCATATCACTGTGATCTTAAACCTTTCGAGGTCTTTGTAGATGCGAACT	2341
QY	2546	CTGCACTCTAAACCCACCTCGCAAGTGCCTTAGCCCGAGCGTACACTTTTCTTAGGC	2605

Db 2342 CTGCACCTCTAAACCCGCTCTGAAGTGCTCTAGCCAGCGCTACACTTTCTTCTAGGC 2401

Qy 2606 ACAATG 2611

Db 2402 NCAATG 2407

RESULT 9

LOCUS BD185304

DEFINITION Helicobacter felis vaccine.

ACCESSION BD185304

VERSION BD185304.1 GI:31877504

KEYWORDS JP 2002355054-A/3.

SOURCE Helicobacter felis

ORGANISM Helicobacter felis

REFERENCE 1 (bases 1 to 2183)

AUTHORS Kusters,J.G. and Cattoli,G.

TITLE Helicobacter felis vaccine

JOURNAL Patent: JP 2002355054-A 3 10-DEC-2002;

AKZO NOBEL NV

COMMENT OS Helicobacter felis

EN JP 2002355054-A/3

PD 10-DEC-2002

PF 16-JUL-2001 JP 2001214711

PI 17-JUL-2000 EP 00202565.8

PR JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI

PC

C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC

A61K39/23,

PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,

PC C12N1/19,

PC

C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC

53,

PC

G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1:01),(C12Q1/68, PC

C12R1:01),

PC C12N15/00,C12N5/00,A61K37/02

CC Helicobacter felis vaccine

Key Location/Qualifiers

FT CDS (3)..(683)

FT CDS (694)..(2181).

FEATURES

source

1..2183

/organism="Helicobacter felis"

/mol_type="genomic DNA"

/db_xref="taxon:214"

ORIGIN

Query Match 67.6%; Score 1948; DB 6; Length 2183;

Best Local Similarity 93.3%; Pred. No. 0;

Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 206 GTGAACCTACACCCAAAGAGCAAGAAAGTCTTGTTATATTATGCGGCGAAGTGGCT 265

Db 3 GTGAACCTACACCCAAAGAGCAAGAAAGTCTTGTTATATTATGCGGCGAAGTGGCT 62

Qy 266 AGAAGCGCAAGACGAGGCGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325

Db 63 AGAAGCGCAAGACGAGGCGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT 122

Qy 326 GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGCCCAAGCTTATGGAAGAG 385

Db 123 GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGCTGAACCTTATGGAAGAA 182

Qy 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGTGGTATATATGTTCCCGAT 445

Db 183 TGTATGCACCTTTTGAAGAAAGATGAAGTATGCCCGGTGGGGAATATGGTCCCTGAT 242

Qy 446 CTAGGTGTAGAAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAATTGGCCCATC 505

Db 243 TTGGGCGTAGAAGCCACTTTCCCGGATGGCACCAAACTCGTAACCGTAATTGGCCCAATT 302

Qy 506 GAACCAAGATGAGCACTTCAAAAGCGGCGAAGTGAATTTGGTTGCGTAAAGACATCGAG 565

Db 303 GAACCTGATGACACACTTTAAAGCGGCGTGAAGTGAATTTGGCTGTGATAAAGACATTGAG 362

Qy 566 CTCAATGCGAGGCAAGGTAACCGAACTTTGAGGTTACTAATGAAGGCGCTTAAATCTCTTG 625

Db 363 CTCAACGTTGGGTAAAGAACTTACCGAGCTTGAAGTTACCAACGAAGGACCTTAAATCTCTTG 422

Qy 626 CATGTGGGTAGCCATTTCCACTTTTGAAGCTAAACAAGCACTTAAATTCGATCGGTGAA 685

Db 423 CATGTGGGTAGCCATTTCCACTTTTGAACCAACAAGGCACTTGAATTCGATCGGTGAA 482

Qy 686 AAAGCCCTATGGCAAAAGCCCTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGA 745

Db 483 AAAGCCCTATGGCAAAAGCCCTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGA 542

Qy 746 CAAACCCGCAAAAGTGCAGTTGATTTCTCTTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC 805

Db 543 CAAACCCGTAAGTGCAGTTTAAATCCCTCTTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAC 602

Qy 806 GGGCTTGTGAATTAACATCGCGGATGAACGCCATATAACATAAAGCGCTTGCAAGGCGAAA 865

Db 603 GGGCTTGTGAATTAATTTGCGGACGAAAGCCCATATAACCAAGGCACTAGACAAAGCAAAA 662

Qy 866 TCTCACGGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAAATATGTAAA 925

Db 663 TCTCACGGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAGTATGTAAA 722

Qy 926 TACTTACGGACCCCAACCAAGGCGATAAAGTGCCTTAGGAGATACCGATTTTGGGCGAGA 985

Db 723 CACCTACGGACCCCAACCAAGGCGATAAAGTGCCTTAGGAGATACCGATTTTGGGCGAGA 782

Qy 986 AGTAGAACATGACTATACCACTTATGGCGAAGAACTTTAAATTTGGCGGGGTAAAACTAT 1045

Db 783 AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT 842

Qy 1046 CCGTAGGGGTATGGTCAGAGCAATAGCCCTGTATGAAAAACACCCCTAGATTTAGTTCATCAC 1105

Db 843 CCGTAGGGGTATGGTCAGAGCAATAGCCCATGAAAAACACCTTAGATTTAGTGTATCAC 902

Qy 1106 TAACGCGGATGATTATCGACTTACACCGGGATTTTCAAAAGCCGACATTTGGGATTAATAACGG 1165

Db 903 CAACGCGATGATTATCGACTTACACCGGATTTTATAAGCCGACATTTGTTATTAATAATGG 962

Qy 1166 CAAAATCCATGGCATTTGGCAAGCGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCA 1225

Db 963 CAAAATCCATGGCATTTGGCAAGCGGGAACAAAGACATGCAAGATGGCGTAAGCCCTCA 1022

Qy 1226 TATGGTCGTGGGTGGGCGACAGAGCACTAGCAGGGGAGGTATGATTATACCGCTGG 1285

Db 1023 TATGGTCGTGGGTGGGCGACAGAGCACTAGCAGGGGAGGTATGATTATACCGCTGG 1082

Qy 1286 GGGAAATCGATTTCACACACCCCACTTCCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345

Db 1083 GGGGATCGATTTCGACACACCCCACTTCCCTCTCCCAACATTTCCCTACCGCTCTAGCCAA 1142

Qy 1346 TGGCGTTTCAACCAATGTTTGGAGCGGCA CAGGTCCTGTAGATGGCAAGTTCGACTAC 1405

Db 1143 TGGTGTGTACCAACCATTTTGGAGGTGGCACAGGTCGCTAGATGGCAAGTTCGACCAC 1202

Qy 1406 TATCACTCCGGGCAATGGAACTTGCACCGCATGTTGGCGCGACGAGAGTATTTCTAT 1465

Db 1203 CATCACTCCGGGCAATGGAACTTGCACCGCATGTTGGCGCGAGCTGAAGAGTATTTCTAT 1262

Qy 1466 GAATGTGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAACT 1525

Db 1263 GAATGTAGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAACT 1322

Qy 1526 AGAAGCGGCGCGATTTGGTTTTAAATTCATGAAGACTTGGGCGCAACAACCAAGTGCAT 1585

Db	1323	AGAACGGGGCGGCGATTGGCTTTAAATTCATGTAAGACTGGGGGCACAAACACCAAGTGGCAT	1382
Qy	1586	CGATCACTGCTTGAGCGTGGCAGATGAATAACGATGTGCAAGTTTGTATCCACACCGGATAC	1645
Db	1383	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATAC	1442
Qy	1646	AGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGGCATCCA	1705
Db	1443	GGTCAATGAGGCGAGGTTATGTGATGACACCCCTAAATGCAATGAACGGGGCGGCATCCA	1502
Qy	1706	TGCCTACCAACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGG	1765
Db	1503	TGCCTACCAACATTGAGGAGCGGGGAGAGCACTCACCTGATGTTATCACCATGGCAGG	1562
Qy	1766	CGAGCTCAATATTTCTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT	1825
Db	1563	CGAGCTCAATATTTCTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT	1622
Qy	1826	TGCAGAACCTTTAGACATGCTCATGACATGCGCACCACTAGACAAACCGCATCCGCGAGGA	1885
Db	1623	TGCAGAACCACTTAGACATGCTCATGACCTGCCACCACTAGATAAGCGCATCCGCGAGGA	1682
Qy	1886	TTTACAAATTTCTCAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTCTCCA	1945
Db	1683	TTTACAAATTTCTCAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTCTCCA	1742
Qy	1946	TGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTCAGCGGA	2005
Db	1743	TGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGCTGGGGA	1802
Qy	2006	AGTGATTCCTCGAACTTGGCAGACTCGCGATAGAAATAAAAAAGAAATTTGGTAAGCTTCC	2065
Db	1803	AGTGATTCCTCGAACTTGGCAGACTCGCAGATAGAAATAAAAAAGAAATTTGGTAAGCTTCC	1862
Qy	2066	TGAAGATGCAAGATAAGATTAATTTCCGATTAAGCGCTATCTCCAAATACACTAT	2125
Db	1863	TGAAGATGCTGAGATAAGCAACATTCGCGCATCAACGCTATATCTCCAAATACACCAT	1922
Qy	2126	CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT	2185
Db	1923	TAATCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT	1982
Qy	2186	CGCCGACTTGGTGTGGAATCTCGCTTTTGGCGTAAAAACCCAAATTCGTGATCAA	2245
Db	1983	CGCCGACTTGGTGTGGAATCTCGCTTTTGGCGTAAAAACCTTAAGATTGTGATTAA	2042
Qy	2246	AGCGGTATGTGGTCTTCTCTGAATGGGCGAATCTAACCGGTCTGTGCCCACTCCCCA	2305
Db	2043	AGGTGGCATGTGGTCTTCTCTGAATGGGCGAATCTAACCGGTCTCGTCCCACTGCTCA	2102
Qy	2306	ACGGTATTATACCGGGAATGTTGGGCATCAACGCAAGGCGGAATTTGACACACGAGAT	2365
Db	2103	GCCGGTTTATACCGGGAATGTTGGGCACCAACGCAAGGCGGAATTTGACACACGAGAT	2162
Qy	2366	CACCTTTGTTTCCAAAGTCG 2385	
Db	2163	CACCTTTGTTTCTCAAGCG 2182	
RESULT 10			
AX356689			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
Helicobacteraceae; Helicobacter.			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL	Patent: EP 1176192-A 7 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1. .2183
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	3. .683
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23259.1"
	/db_xref="GI:18674027"
	/translation="MKELTPKQEKFLYYAGEVARKKAEGLKLNQNPALAIYISAHIM DEARRGKTVAAELMECWFLKKDVMVPGNMVVDLGEATFPDGTGLKLVNWNIEP DEHPKAGEVFKGDKDIELNVKGEVTELVNEGPKSLHVGSHFHFPTNFKALKFDR KAYGKLDIPSGNLTIRIGAGTRKVLQILPGGSKKVIQMNGLVNNIADERHKHKALDK AKSHGFIK"
	694. .>2181
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23260.1"
	/db_xref="GI:18674028"
	/translation="MKMKKQBYNVTGPTTGDKVRLGDTDLWAEVEHDDYTTYGBELKF GAGKTIREGMSQNSPDENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKANGD MODGVKPMVVGVTALAGEMIIITAGGIDSHTHFLSPQOPFLKANGHITGFGGTT GPVDGTNATTITPKGNLHRLRAABEYVMNVGFLGKGNSSKKQLVBOVBAGATGFK LHEDWGTTPSAIDHCLSVADBYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAVHIEG AGGSHSPDVTIMAGELNLPSSYPTTPIYTINTVAEHLMDLMTCHLLDKRLREDLQFS QSRIRFGSIAAEDVLDHIGVIAMTSSDSQAMRAGEVIPTWQTADRNKKBFKULPED GADNDFRIKRYISKYTINPALTHVSEYIGSVEEGKIADLVVWNPAPFFGVKPKIVIK CGMVVFSEMGSDNASVPTQPVPVYVREMFHGHGKAKFDTSITFRVSS"
ORIGIN	
	Query Match 67.6%; Score 1948; DB 6; Length 2183; Best Local Similarity 93.3%; Pred. No. 0; Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
Qy	206 GTCAAACTCACCCCAAGAGCAGAGGCTTAAAGCTCAACCAACCCGAGCAAGTGGCT 265
Db	3 GTCAAACTCACCCCAAGAGCAGAGGCTTAAAGCTCAATCAACCCGAGCAAGTGGCT 62
Qy	266 AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAGCAAGTGGCTTACATTAGT 325
Db	63 AGAAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGAGCAAGTGGCTTACATTAGT 122
Qy	326 GCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTGTCGCCAGCTTATGGAAGAG 385
Db	123 GCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTGTCGTAACCTTATGGAAGAA 182
Qy	386 TGCATGCACTTTTGAATAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 445
Db	183 TGTATGCACTTTTGAATAAAGATGAGTGTATGCCGGGTGGGGAAATATGTTCCCTGAT 242
Qy	446 CTAGGTGTAGAAGCCACCTTTTCCCTGATGCTACGAAACTTGTAACTGTGAAATTTGCCCATC 505
Db	243 TTGGGCGTAGAAGCCACTTTCCCGATGGACCAACTCGTAAACCGTGAATTTGGCCCATT 302
Qy	506 GAAACAGATGAGCACTTCAAAGCGGCGGAAAGTGAATTTGGTTCGCTAAAGACATCGAG 565
Db	303 GAAACCTGATGAACACTTTTAAAGCGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTGAG 362
Qy	566 CTCATGTCAGGCAAGAGTAGTAAACCGAATTCAGGTTACTAATGAAGGCCCTAAATCCTTG 625
Db	363 CTCACGTGGGTAGGAAGTTACCGAGCTTTGAAGTTTACCAACGAAGGACCTTAAATCCTTG 422
Qy	626 CATGTGGGTAGCCATTTCCCACTTCTTTGAAGCTAACCAAGGCACTAAATTCGATCGTAAA 685
Db	423 CATGTGGGTAGCCATTTCCCACTTCTTTGAAACCAACCAAGGCAATTTGAAATTCGATCGGAAA 482
Qy	686 AAAGCCTATGGCAAAACGCTTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 745

Db 483 AAAGCCTATGGCAACGCCCTAGATATTCCTCTGGCAACACGCTACGCATTTGGGGCAGGA 542
Qy 746 CAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATGGCATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTTAATCCTCTTGGCGGTAGTAAAAAAGTGAATGGCATGAAC 602
Qy 806 GGGCTTTGTGAATAACATCGCGGATGAACCGCATAAACATAAAGCGCTTGAACAAGGGGAAA 865
Db 603 GGGCTTTGTGAATAATATTCGGGACGAACCGCATAAACATAAAGCGCATTAAGCAAGGGGAAA 662
Qy 866 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATAATGAAAAAACAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATAATGAAAAAACAAGAGTATGTAAA 722
Qy 926 TACCTACGGACCCACCAAGCGCATAAAGTGCCTTAGGAGATACCGGATCTTTGGGCAGA 985
Db 723 CACCTACGGACCCACCAAGCGCATAAAGTGCCTTAGGAGATACCGGATCTTTGGGCAGA 782
Qy 986 AGTAGAACATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 783 AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAAAATTTGGCGCGGTAAAACTAT 842
Qy 1046 CCGTGAAGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACACCTAGATTAGTCACTAC 1105
Db 843 CCGTGAAGGTATGGGTACAGAGCAATAGCCAGATGAAAAACACCTTAGATTAGTCACTAC 902
Qy 1106 TAAACCGCATGATTATACCTACACCGGGATTTAAAGCGCAGATTTGGGATTTAAAAACGG 1165
Db 903 CAACCGCATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGGTATTAAAAATGG 962
Qy 1166 CAAATTCATGGCATTTGGCAAGGAGGAAACAGAGCATGCAAGATGGCGGTAGCCCTCA 1225
Db 963 CAAATTCATGGCATTTGGCAAGGAGGAAACAGAGCATGCAAGATGGCGGTAGCCCTCA 1022
Qy 1226 TATGGTCTGGGTGGTGGCAGAGCACTAGCAGGGGAGGTATGATTATTCACCGCTGG 1285
Db 1023 TATGGTCTGGGTGGTGGCAGAGCACTAGCAGGGGAGGTATGATTATTCACCGCTGG 1082
Qy 1286 GGGAAATCGATTTCACACACCCACTTCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
Db 1083 GGGGATCGATTTCGACACACCACTTCCTTCTCCCACTTCCCTACCGCTCTAGCCAA 1142
Qy 1346 TGGCGTTAACACCATGTTTGGAGCGGCA CAGGTCCTGTAGATGGCAGAAATCGCATAC 1405
Db 1143 TGGTGTGTACACCATGTTTGGAGTGGCAGAGTCCGGTAGATGGCAGAAATCGCATAC 1202
Qy 1406 TATCACTCCGGCAAAATGGAATCTGCACCGCATGTTGGCGGCAGCAGAGATTTCTAT 1465
Db 1203 CATCACTCCGGCAAAATGGAATCTGCACCGCATGTTGGCGGCAGCTGAAGAGTATTCAT 1262
Qy 1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAAAT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAAAT 1322
Qy 1526 AGAAGCGGGCGGATTTGGTTTAAATTCATGAGACTGGGGCAACACCAAGTGGGAT 1585
Db 1323 AGAAGCGGGCGGATTTGGTTTAAATTCATGAGACTGGGGCAACACCAAGTGGGAT 1382
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGCATAC 1645
Db 1383 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGCATAC 1442
Qy 1646 AGTCAATGAGCGAGTTATGATGATGACACCTAAATGCAATGAACGGGCGGCATCCA 1705
Db 1443 GGTCAATGAGCGAGTTATGATGATGACACCTAAATGCAATGAACGGGCGGCATCCA 1502
Qy 1706 TGCCTACCATTTGAGGAGCGGTGGAGGACACTCACCTGATGTATTCACCATGGCAGG 1765
Db 1503 TGCCTACCATTTGAGGAGCGGTGGAGGACACTCACCTGATGTATTCACCATGGCAGG 1562
Qy 1766 CGAGCTCAATTTCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATTTCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1622

Qy 1826 TGCAGAACACTTAGACATGCTCATGATGCCACCACTTAGACAAAACGCAATCCCGCAGGA 1885
Db 1623 TGCAGAACACTTAGACATGCTCATGATGCCACCACTTAGATAAGCGCATCCCGCAGGA 1682
Qy 1886 TTTTACAAATTTCTCAAGCCGATTCGCGCCCGGCTCTATCGCGCTCAAGATGTGCTCCA 1945
Db 1683 TTTTACAAATTTCTCAAGCCGATTCGCGCCCGGATCTATTTGCGCTGAGGATGTGCTCCA 1742
Qy 1946 TGATATGGGTGTGATTCGCGATGCAAGCTCGGATTCGCAAGCAATGGGCGCTGCAGGCGA 2005
Db 1743 TGATATTTGGGTGTGATTCGCGATGCTAGCTCCGATTCGCAAGCAATGGGCGCTGCAGG 1802
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAGAAATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1803 AGTGATTTCTCGAACTTTGGCAGAACTGCAGATAAGATAAAAAAGAAATTTGGTAAGCTTCC 1862
Qy 2066 TGAAGATGGCAAGACATAACGATAATTTCCGCAATTAAGCGGTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGGGTGAGATAACGACAACTTCGCGATCAAAACGCTATATCTCCAAATACACCAT 1922
Qy 2126 CAAACCCCGCTTTGACCCACGCGTGAAGTATATCGGCTCTGTGGAAGAGGCAAGAT 2185
Db 1923 TAATCCGCTTTGACCCATGCGGTGAGGAGTATATCGGCTCTGTGGAAGAGGCAAGAT 1982
Qy 2186 CGCGCACTTGGTGGTGGAACTCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1983 CGCGCACTTGGTGGTGGAACTCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2042
Qy 2246 AGCGGTATGGTGGTCTCTCTGAAATGGGAGTCTAAACGCGTCTGTGCCCATCTCCCA 2305
Db 2043 AGTGCGCATGGTGGTCTCTCTGAAATGGGAGTCTTAACGCGCTCCGTCGCCACGCTCA 2102
Qy 2306 ACCGTTTTATTACCGCAAAATTTTGGGATCAAGCAAGCGGAAATTTGACACCAAGCAT 2365
Db 2103 GCGGTTTTATTACCGCAAAATTTTGGGATCAAGCAAGCGGAAATTTGACACCAAGCAT 2162
Qy 2366 CACTTTTTTTTCCAAAGTCG 2385
Db 2163 CACTTTTTCGTGTCTCAAGCG 2182

RESULT 11
AF330621 8406 bp DNA linear BCT 13-MAY-2003
LOCUS Helicobacter bizzerononii tRNA ribosyltransferase gene, partial
DEFINITION cds; GidB, complete cds; urease gene cluster, complete sequence and
unknown gene.
ACCESSION AF330621
VERSION AF330621.1 GI:27462193
KEYWORDS
SOURCE Helicobacter bizzerononii
ORGANISM Helicobacter bizzerononii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 8406)
AUTHORS Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C.,
McDonough, P., McDonough, S. and Chang, Y.F.
TITLE Cloning and characterization of a Helicobacter bizzerononii urease
gene cluster
JOURNAL DNA Seq. 13 (6), 321-331 (2002)
PUBMED 12652903
REFERENCE 2 (bases 1 to 8406)
AUTHORS Zhu, J. and Chang, Y.F.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) Population Medicine and Diagnostic Science,
College of Veterinary Medicine, Cornell University, Tower Road,
Ithaca, NY 14853, USA
FEATURES
source
1..8406
/organism="Helicobacter bizzerononii"
/mol_type="genomic DNA"
/db_xref="taxon:56877"

HECUREBASE LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM

HECUREBASE 2664 bp DNA linear BCT 03-MAY-1995
 Helicobacter heilmannii urease, complete cds's.

L25079 L25079.1 GI:793908
 urease.

Candidatus Helicobacter heilmannii
 Candidatus Helicobacter heilmannii
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 Helicobacteraceae; Helicobacter.

1 (bases 1 to 2664)
 Solnick, J.V., O'Rourke, J., Lee, A. and Tompkins, L.S.
 Molecular analysis of urease genes from a newly identified
 uncultured species of Helicobacter
 Infect. Immun. 62 (5), 1631-1638 (1994)

JOURNAL PUBLISHED 8168924
 COMMENT On May 3, 1995 this sequence version replaced gi:529422.
 Original source text: Helicobacter heilmannii (individual isolate
 2) DNA.

FEATURES source

Location/Qualifiers
 1..2664
 /organism="Candidatus Helicobacter heilmannii"
 /mol_type="genomic DNA"
 /isolate="2"
 /db_xref="taxon:35817"
 /lab_host="BALB/c"
 200..204
 211..915
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65722.1"
 /db_xref="GI:529423"
 /translation="MKLTPELKDMLHVAGELAKORKAKGIKLNTVEAVALISAHV
 EARAGKSVADIMQSGRTLLKDDVMPGVAHMHHVGVGEAFPDGTGLVTHPTPEVA
 GSKLAPGVILKNDITLNAHGKHAVOLKNGKDRPQVGHSHFVFEVKNLDDFORE
 KATGKRLDIASGTAVFPEPGEETKLTIDIGNKRIYGFNALVDROADHDKKLAUKR
 AKEKHPTGTCNGCDNR"

RBS 919..923
 CDS 930..2636
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65723.1"
 /db_xref="GI:529424"
 /translation="MKKISRKEYSVNMGPTTGKVKVLGDTLDLILEVEHDCTTYGEBIK
 FGGKTRIDMGMTGNSSPSHSLDVLITNALIVDTGYIKADIGIKNGKIHGKAGNK
 DLQDGVNRLCVGPATEALAAEGLITAGGIDITHIFI SPOQIPTAFASGITTMIGGG
 TPGADGTNATITTPGRNNILKEMLRASEYAMNGLYKGVNFPFEPALIDQLEAGATGF
 KIHEDGSPSIAVFNALINADKYDVQVAIHTDITLNEAGCVEDTLEAIAGRTIHTPTE
 GAGGHGAPDVKWAGENLITPASTNPTIPTKNTAEHMDMLCHLDKNIKEDVFA
 DSRIRQTTAAEDKLDHMGIFSTSSDSQAMGRVGVITRTQTADKNKKGFGRLPEE
 KGDNDNFRIKRYISKYITINPAITHGISSEYGVSGVKYADLVLSPAFFGIGKPNMIIK
 GFIALSQMGDANASITPTQPVYIRFMFGHGKAKEDTNI TFSQVAYENGIKHLEIG
 QRVLVPVQNCRNRTKKDLKENDVTAHIEVNPETYKVKVDGNEVTSHAADKLSLAQLYN
 LF"

RBS
 CDS

ORIGIN

Query Match 39.5%; Score 1138; DB 1; Length 2664;
 Best Local Similarity 66.4%; Pred. No. 6.8e-217;
 Matches 1709; Conservative 1; Mismatches 826; Indels 36; Gaps 4;

QY 63 TTGCTTAATRCATATTAATTTTTTAAATTAATCTACTTATTATCATATATATATATATT 122
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 70 TTGTGAATAGATTAACAAAGAAATAATAAT--ATTATCTTTGGATAACAAATATTAC 127
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 123 ACTTATTATTAATAAAGTTAATAAAGTAACGAAATAGGACTATAATCCCATTCGCTTTA 182
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 128 AACCATTAGCAAAAGACTAATAAACTGCTCGTTGATAGCTTGCCTATTCAATAACA 187
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 183 AAATTTAACAACAGGAGTAAATAGGTGAACCTCACCCAAAGACGAAGAAAGTCTTTGT 242
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 188 CCATTTTATACGAGGAGTTTAGATGAACCTGACCTTAAGAGTTGGATAAGTTCATGC 247
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1292 CGATTACACACCCACTTCTTCTTCCCAAAAGATTCCTTCCGCTCTAGCCATGCGGT 1351
Db 1328 CGACACCCACATCCACTTCAATTTCTCTCAACAAATCCCTACTGCTTTTGTAGCGGAAT 1387
QY 1352 TACAACCATGTTTGGAGGGGACAGTCTCTGTAGATGGCAGCAATGCGACTACTATCAC 1411
Db 1388 CACCACCATGATCGGGGGGGGACTGAGCTGCGATGSCACCAACGCCACCAATCAC 1447
QY 1412 TCGGGCAAAATGGAATTCGACCGCATGTTGCGCGCAGCAGAGATGATTTCTATGAATGT 1471
Db 1448 TCCGGCGCGTGGAACTTAAAGAAATGCTCCGCGCTTCTGAAGAAATACGCCATGAACCT 1507
QY 1472 GGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAACAACTTGTGTAGAACAGTAGAGC 1531
Db 1508 TGCTTACCTTGGTAAGGGGAATGTCTTTTGAACCTGCTCTCATTTGACCGAGCTAGAAGC 1567
QY 1532 GGGCGGATTTGGTTTAAATTTGATGAAGACTGGGGCACAACACCAAGTGCATCGATCA 1591
Db 1568 CGGCGGATTTGGTTTAAATTTGATGAAGACTGGGGCAGCAGCACTTCCAGCATCAACCA 1627
QY 1592 CTGCTTTGAGCGTGGCAGATGAATPACGATGTGCAAGTTTGTATCCACCGATACAGTCAA 1651
Db 1628 CGCTCTAAACATCGCTGACAAATACGATGTGCAAGTGGGATCCACACGACACCTTGAA 1687
QY 1652 TGAGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGCGCGCCATCCATGCTTA 1711
Db 1688 TGAAGCGGCTGTGTGGAAGACACCTTAGAAGCCATCGCTGGACGACCACTCCACACCTT 1747
QY 1712 CCATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGACGCGAGCT 1771
Db 1748 CACACGGAAGGTGTGGCGGCGACACGCTCTGTGAGTGAATGCGGCGGCAAT 1807
QY 1772 CAATATTTACCTCTCTCCACACCCCACTATTTCCCTTATPACCATTAATACGGTTGCGA 1831
Db 1808 TAACATCTTCTCTCTTACCAACCCCACTATTTCCACCAAAAACACAGAAGCGA 1867
QY 1832 ACATTTAGACATGCTATGATGACACCTGACACCTAGACAAACGATCGCGAGATTTACA 1891
Db 1868 ACATGGAATGCTTTATG---TGCCACCACTTTGGATAAACCAATCAAGAGATGGGA 1924
QY 1892 ATTTTCTCAAAGCGTATCGCGCCCGCTCTATCGGGCTGGAAGATGTGCTCCATGATAT 1951
Db 1925 ATTGCGGACTACGATATCGCGCCCAACCATTTGGCGGGAAGATTAATCCACGACAT 1984
QY 1952 GGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGAGCGCAAGTGAT 2011
Db 1985 GGGATTTTCTCATCTACTAGTCCGACTCTCAAGCCATGGTGTGCGTGGCGGAGTGAT 2044
QY 2012 TCCTGCAACTTGGCAGACTCGGATAGAAATAAAGAAATTTGGTAAGCTTCTCTGAAGA 2071
Db 2045 CACCCGCACTTGGCAACAGCGGACAAAACAAAAGAAATTTGGCGCTTGGCTGAAGA 2104
QY 2072 TGCAAGAATAACGATTAATTTCCGATTAAGCTACATCTCCAAATACATATCAACCC 2131
Db 2105 AAAAGGCGACAAAGCAACTTCGCAATCAACCGCTATTTTCCAAATACACCATCAACCC 2164
QY 2132 CGCTTTGACCCACCGCTGAGCAGATATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGGA 2191
Db 2165 TGCATCACACACGCGATTTCTGNATATGTGCGCTCTGTAGAGTGGGTAAATACGCTGA 2224
QY 2192 CTTTGGTGTGTGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGTATCAAGGCGG 2251
Db 2225 CTTTGGTGTGTGGAGCCCTGCTTCTTTGGCATTAACCCCAACATGATATCAAGGCGG 2284
QY 2252 TATGTTGGTCTTCTGTAATGGCGATTTCTAAGCGGTCTGTGCCCACTCCCAACCGGT 2311
Db 2285 TTTTATTCGCGCTTTCTCAAAATGGCGATGCGCAACGCTTCTATCCCACTCCCAACCGGT 2344
QY 2312 TTATTTACCGGAATGTTTGGGATCACGCAAGCGGCAAAATTTGACACCAAGCATCACTTT 2371
Db 2345 GTACTACCGGAATGTTTGGGCAACCGGCAAGCGCAAAATTTGACACCAAGCATCACTTT 2404
QY 2372 TGTTTCCAAAGTCGCTATGAAAATGGCGTGAAGAAAAGCTGGGCTTTAGAGCGCCCAAGT 2431

Db 2405 CGGTCTCAAGTGGCTTATGAAAACGGCATTAACACGAGTTGGCTTGCAAAAGTTGT 2464
QY 2432 TCTACCGGTCAAAAACATGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAAC 2491
Db 2465 GTTGCCTGTGAAAACATGCGCGCAACATCACCAAAAAGACCTTAAGTTCAACGATGTAC 2524
QY 2492 GGCAAAATCACCGTGCATCGGAAACCTTCGAGGTCTTTGTAGATGCGAACTCTGCAC 2551
Db 2525 CGCACACATCGAAGTCAACCTGAAACCTCAAAAGTGAAGTGGATGCAACGAGTTAC 2584
QY 2552 CTCTAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAG 2603
Db 2585 CTCCATGCGGCTGACAAATTTGAGCCTGACCAACTCTATACCTATTTCTAG 2636

RESULT 13
AE001446/c 12037 bp DNA linear BCT 20-JAN-1999
LOCUS Helicobacter pylori, strain J99 section 7 of 132 of the complete
DEFINITION genome.
ACCESSION AE001446 AE001439
VERSION AE001446.1 GI:4154573
KEYWORDS
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 12037)
AUTHORS Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C.,
Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G.,
Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C.,
Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E.,
Vovis, G.F. and Trust, T.J.
TITLE Genomic-sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori
NATURE 397 (6715), 176-180 (1999)
9923682
2 (bases 1 to 12037)
AUTHORS King, B.L., Alm, R.A. and Trust, T.J.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney
Street, Cambridge, MA 02139, USA
COMMENT Address all correspondence to: hp@arcb.us.astro.com or Richard
A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge,
MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith,
Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills,
Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics
Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and
Diane E. Taylor are with the University of Alberta Department of
Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G
2H7 and the Canadian Bacterial Diseases Network. All other authors
are with Astra Research Center Boston, 128 Sidney Street,
Cambridge, MA, 02139. Putative identifications, sequence
alignments, and name and sequence search capability are available
at ARCB's World Wide Web site. (URL:
http://www.astra-boston.com/hpylori).
FEATURES
source
1..12037
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"
/db_xref="taxon:85963"
complement(143..1852)
/gene="ureB"
/note="synonym: jhp0067"
complement(143..1852)
/gene="ureB"
/note="similar to H. pylori 26695 gene HP0072"
/codon_start=1
/transl_table=11
/product="UREASE BETA SUBUNIT"
/protein_id="AAD05651.1"

```

/db_xref="GI:4154576"
/translation="MKKTSRKEYVSMYGGTTGDKVRLGSDPTDLIAEVEHDDYTYIGBELK
FGGKTLREGMSQSNPSKEELDLIITNALIVDYTCIYKADIGICKGKIAGIKGKGNK
TMDQGVKNLSVGPATELAGELIIVTAGCIDTHIFISPOQIPTAFASGVTTMIGGG
TPADGTWATITTPGRNLKWLRAAEYSMNLGFLAGNASNDASLADQIEAGAI GF
KIHEDWGTTPSAINHALDVADKYDQVAIHDTDLNEAGCVEDTMAAIAGRTWHTFTE
GAGGHHADPIIAGEHNILPASTNPTTPTFTVTEAEHMDMLVMVCHHLDKSLKEDVQF
ADSRPTQIAEDTLHDMGIFSIITSSDQAMGRGVEITRTWQADKNKKEFGRLKE
EKGDNDFRIKRYLTPAIAGISYVGVSVGVKADLVLSWSPAFGVGPKNMII
KGGFTALSQMGDANASITPQPVYVREMAFHGKAKYDANITVFSQAAVDKGIKEELG
LEROVLVYKNCNITKDMQFNDTTHAIEVNPETYHVFDVKGVTSPKANKYSLAQLF
SIF"
complement(1856..2572)
/gene="urea"
/note="synonym: jhp0068"
CDS
complement(1856..2572)
/gene="urea"
/note="similar to H. pylori 26695 gene HP0073"
/codon_start=1
/transl_table=11
/product="UREASE ALPHA SUBUNIT"
/protein_id="AAD05652.1"
/db_xref="GI:4154577"
/translation="MKLTPEKLDKMLHYAGELARRKKEGKILNYVEVALISAHIM
BEARAGKTAELMQEGRLLKPDDDMDGVASMIHEVGIEAMPDPGTLVTVHTIEA
NGKLVGELFLKNEIDITINEGKKAHSVKNVGNDRPVOIGSHFHFVFNRCIDFDEK
TFGRDLIASGTAVRFEPEEKEKVELIDIGNRRIFGFNALVDRQADNESKKIALHRA
KERGFHGAKSDDNYVYKIE"
complement(2926..3417)
/gene="lspA"
/note="synonym: jhp0069"
CDS
complement(2926..3417)
/gene="lspA"
/note="similar to H. pylori 26695 gene HP0074"
/codon_start=1
/transl_table=11
/product="LIPOPROTEIN SIGNAL PEPTIDASE"
/protein_id="AAD05653.1"
/db_xref="GI:4154578"
/translation="MLNTQKSLLVPMGVFFLIFGVDAQIAKHAILLEGHYBSLVLDIV
LVFNKGAFSLFLEGGELKYLQILLIGLFIPLMRQKELFKSHAIEFGMVFGVSN
LDRFVGGGVDDYVYHYGFDFPAIFNFADVMIDVGVGVLLLRQFFFKQNKIRGIIT
LPK"
complement(3411..4748)
/gene="glmM"
/note="synonym: jhp0070"
CDS
complement(3411..4748)
/gene="glmM"
/note="similar to H. pylori 26695 gene HP0075"
/codon_start=1
/transl_table=11
/product="phosphoglucosamine mutase"
/protein_id="AAD05654.1"
/db_xref="GI:4154579"
/translation="MKIFGTGVRGKAGVKLTPEMVRGLGIAAGLYFKKHSOTNKILL
GKDTKSCMYENALVSALTSGYNNVIOIGPMPPTAIAPLFTEDMCDAGIMISASHNP
FEDNGIKFNSYGYKKEEBERAIEI FHDSELLSSYKVGSVSASAKRIDDIVGRYI
AHLKHSFPKHLNLQKLRIVLDTANGAAYKAVPVFSELGADVLVINDPENGINNEQC
GALHPNQLSVKRYTRADLGFADGADRLVVDNLGNI VHGDKLLGVLYVYQSKNA
LSQAVATNSMLALKEYLKQDLKELKCAIGDFVSECMRLKNAIFGQSGSHIIF
SDYAKTGDLGVALQSVLLESKLVSIALNPFLYPQSLVNLNIOKKPPLSESUGY
SALLKELDLQELIRHLIRYSGTENKLRILLEAKDEKLESKQELKEFFEGQLC"
4838..5107
/gene="rpst"
/note="synonym: jhp0071"
CDS
4838..5107
/gene="rpst"
/note="similar to H. pylori 26695 gene HP0076"
/codon_start=1
/transl_table=11
/product="30S RIBOSOMAL PROTEIN S20"
/protein_id="AAD05655.1"
/db_xref="GI:4154580"

```

```

/gene
CDS
/translation="MANHKSABKRIQTIKRTERNRFYTKVKNIIKAVREAVAINDV
AAQSRRLTANKELHKFVSGKILKKNATSRKVSRLNASVKKIALA"
5226..6284
/gene="prfA"
/note="synonym: jhp0072"
CDS
5226..6284
/gene="prfA"
/note="similar to H. pylori 26695 gene HP0077"
/codon_start=1
/transl_table=11
/product="PEPTIDE CHAIN RELEASE FACTOR 1"
/protein_id="AAD05656.1"
/db_xref="GI:4154581"
/translation="MSILAEKLSILKRYDELTALESSVEVVSIDIKKLTELSKEQSSI
EISVASKEYLSVLEIGIKENKELLEDELSAKELKILEIKSELETAIKOLLIPK
DNDKNIYLELRAGTGGDEAGIFVGDLPKAYCRYADLKKWKEIVISSSENSVGKYKE
IIVLJKGVYSRLKFEAGTRVORVPETESQGRIRHTSAITVAIMEVDVDSINPS
DLKIEVFRAGHGGQCVNTDSAVRITHLPTNISVSMQDEKSHQKDKALKILKARL
YEKQIEEQOLANAKORKEQVSGDRSERIRTYNPQRJUSEHRINLTLSLEIMLSG
NLDVINPLIAHAQSQFE"
6698..7465
/gene="jhp0073"
CDS
6698..7465
/gene="jhp0073"
/note="similar to H. pylori 26695 gene HP0078 and HP0079"
/codon_start=1
/transl_table=11
/product="putative Outer membrane protein"
/protein_id="AAD05657.1"
/db_xref="GI:4154582"
/translation="MKKIFLGMALAFSVSMAEKSGAFLGGGFQYSNLENQNTTRTPSA
NNNTPINTSMFGNNOAPAQETPSVINNNYQMGYGDAMAGYKWFPGTKRFGFRY
GYVSNHANLSFVSGSKLGIIMDGASOVNNFTYGVGFDALYNFYESKGYNTAGL.FVGFG
LGDSFIVQGESYLSQMQICNNTAGCSAMTSYEQMPVEFGFRSNFSKHSGIEVGF
KULPFTNIFYKERGVDGSDVDFYKRNFSIYFNMINL"
7716..9488
/gene="jhp0074"
CDS
7716..9488
/gene="jhp0074"
/note="similar to H. pylori 26695 gene HP0080"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD05649.1"
/db_xref="GI:4154574"
/translation="MKAAILIIMTILSLNAISVNRALFDLQDSQLKGLTPKIVDFGG
YKSNTEWCATALNYINAANGDAKPSALVERKPNFSGILGNFRAHHLKQALKLQKN
LKYLKLI IARDSFYSGTYRTGIYIPLGISLKDQKTAQKMLADLSVVGAYLKQOENKRAQ
SPYRSNNYVNSYSPYGYMGYGMGYMTGMGYDFYDFYDGYGFYNNFMFMQ
VODYLMENYVALDQEEILDHDAS INOLDTPTDDDDKDKSOPANLMSFYRDPK
FSKDIQTNRILNSALVNLNLSHMLKDNSLPHTKAMPTKSDVAITSQAKLNHLVGQIKE
MKQDASPKNKIDSVVNKAMEVRDKLDNLNLQDLKQKLSSEQAQVDAKALDSVQ
QLSHSDVVGNYLDGSLIDGDDRDDLDANINPMQPAQQTPINNMDNTHANDSKDQ
GNGALINPNNAITNDHNDHMDTNTTDSANANDPTTDDKQASGNTGDMNTDTGNTD
TGNTDTGNTDMSNMNGNDGTGNTNDMDGNSNDMGDMNANDMDMGNSNDMDGD
MGDMNDMGDMGDMGDMGDMGN"
complement(9912..11933)
/gene="jhp0075"
CDS
complement(9912..11933)
/gene="jhp0075"
/note="similar to H. pylori 26695 gene HP0082"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein (MCP)"
/protein_id="AAD05650.1"
/db_xref="GI:4154575"
/translation="MKSTRIGSKIVMVCAVVIVISAVMGVVIYSKVSVLQSOATEL
LOKKAQLYSFKIQIGIKELPMGANTLERFLSDENGAINDTLKERMISEFLANPHVLL
VSAIYNNRMITANNMDSKIAFYNTALNENWTNQHLSKSTRSDPYKVEYNGKIL
YGMIDITLPMGKNQNVIGALNFFNLIDAFYTDVVGKKKNTFLALQIPDFPEEGNGN
QDKILSAINPDKRAKAVEYQNQNEAGTSLHSGNTTETFLAIQIPDFPEEGNGN
HWRWAIGYVKNKSLVFEKATNTKFIITTLILGLVLFLVFLFIISNLITKRISRVNN

```

Query Match	37.8%;	Score 1090.8;	DB 1;	Length 12037;
Best Local Similarity	65.3%;	Pred. No. 1.4e-207;		
Matches 1681;	Conservative 1;	Mismatches 853;	Indels 39;	Gaps 4;
QY	67	TAATRCATTATTAAATTTTAAATAATTAATCTT-ATTATCATATATAATAATATTACT	125	
DB	2712	TAATACTAATCAATTTAAATAATAATATTAGTTAATAACGCTTCTGTTAATCTTAGTAAA	2653	
QY	126	TATATTTAAAAAGTTAATAAAAAAGTAACGAAATTTAGGACTATATCCCATTTGCCCTTTAAAA	185	
DB	2652	TCAAAAAGATTGCTFACAATTACTCCAACTTGAATGGTATTGCTTCAAGGAAAAAACAC	2593	
QY	186	TTTAAACAAAGAGTAAATAGGTGAATCTACACCCAAAGCAAGAAAGTTCTTTGTTAT	245	
DB	2592	TTTTAAGAATAGGAGAATGAGATGAATCAACCCAAAAGAGTTAGATAAGTTGATGCTCC	2533	
QY	246	ATTATCGGGCGAAGTGGCTAGAAAGCGGCAAGCAGAGGGCTTAAAGCTCAACCAACCCG	305	
DB	2532	ACTACGCTGAGAAATTAGCTTAGGAACGCAAGAAAGAAAGGCAATTAAGTTAATATGTGG	2473	
QY	306	AAGCCATTGCTTACATTTAGTGCCCATATTTATGGACGAAGCGCGCTGGAAAAAACC	365	
DB	2472	AACGGTAGCTTTGATTAGTGCCCATATTTATGGAAGAGCGAGCTGGTAAAGACTG	2413	
QY	366	TTGCCAGCTTATGGAAGAGTGATGCATCTTTTGAAGAAAGATGAAGTAATGCCCGGG	425	
DB	2412	CGGCTGAATTTGATGCAAGAAAGGCGCACTCTTTTAAACCGGATGATGTGATGGATGGT	2353	
QY	426	TGGGTAAATAGTTTCCGATCTAGGTGTAGAGCCACTTTTCTGATGGTACCAAACTTG	485	
DB	2352	TGCAAGCATGATCCATGAAGTGGTATTAAGCGATGTTTCTGATGGGACCAAACTCG	2293	
QY	486	TAACTGTGAATTTGGCCCATCCAAACCATGAGACACTTCAAAGCGGGCGAAGTGAATTTG	545	
DB	2292	TAAACCTGTCATACCCCTATTAGGCTTAATGTTAAATGGTTCTGTTGAGTTG---TTC	2236	
QY	546	GTTGCGATAAAGACATCGAGCTCAATCGACGGCAAGAAAGTAACGAACTTTGAGGTTACTA	605	
DB	2235	TAAAAAATGAAGACATCACTATCAACGAGGCAAAAAAGCCGTAGCGTGAAGTTAAAA	2176	
QY	606	ATGAAGGCCCTAAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTTGAAGCTTAA	665	
DB	2175	ACGTGGCGCAGACACCGGTTCAAATCGGTTCAACACTTCCATTTCTTTGAAGTGAATAGAT	2116	
QY	666	CACATAAAATTCGATTCGTGAAAAAGCCATATGGCAAAACGCTTAGATATTCCTCTGGCAACA	725	
DB	2115	GCCTAGACTTTGACAGAGAAAAAACTTTCCGCAAAACGCTTAGACATTTGCCAGCGGACAG	2056	
QY	726	CGCTACGCATTTGGGGCAGGACAAACCCGCAAAAGTCAGTTGATTCCTTTTGGTGGCAGTA	785	
DB	2055	CGGTAAAGTTTGAAGCTTTGAGCGAATAACGTTAGATTTGATAGGCAAGCAGACCAAGCAAAAA	1996	
QY	786	AAAAAGTGAATGGCATGAAACGGGCTTGTGAATAACATTCGGGATGAACGCCATAAACATA	845	
DB	1995	GAAGAATCTTTTGGATTTTAAACGATTTGATAGGCAAGCAGACCAAGCAAGCAAAAAA	1936	
QY	846	AAGCGCTTGACAAAGCGAAATCTCAACGATTT-----	877	
DB	1935	TTGCTTTTACAGAGCTTAAGAGCGTGGTTTTTCAATGCTGTAAAGCGATGACAACTATG	1876	
QY	878	----ATCAAGTAAAGGAGCTCCCATGAAA---ATGAAAAAACAGAAATATGTAATACCT	930	
DB	1875	TAAAAACAAATTAAGAGTAAAGAAATGAAGAAATAGCAGAAAAAGATATGTTTCTATGT	1816	
QY	931	ACGACCCCAACAAAGGCGATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAGAAGTAG	990	
DB	1815	ATGGCCCTACACAGGGGATAAAGTGAATTTGGGCGATACAGATTTTGATCGCTGAAGTAG	1756	
QY	991	AACATGACTATACCACCTATGGCGAAGCACTTAAATTTGGCGCGGTAAACATATCCGTG	1050	
DB	1755	AACATGACTACACCATTTATGGCGAAGACTTTAAATTTGGTGGCGGTAAACCCCTAAGAG	1696	
QY	1051	AGGATATGGGTGACAGCAATAGCCCTGATGAAAAACACCCCTAGATTTTATGTCATCACTAACG	1110	
DB	1695	AAGGCATGAGCCAAATCCAAACAAACCCCTAGCAAAAGAAACTGGATTTGATCATCACTAACG	1636	
QY	1111	CGATGATTTACACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACCGCAAAA	1170	
DB	1635	CTTTAATCGTGAATTTACACCGGTATTTATTAAGCGGATATTGGTATTTAAAGACGCGCAAAA	1576	
QY	1171	TCATGCGCATTTGGCAAGCGAGAAACAAAGCATGCAAGATGGCGTAAGCCCTCATATGG	1230	
DB	1575	TCGCTGGCATTTGGTAAAGCGGTAAACAAAGACATGCAAGATGGCGTTAAAAACAATCTTTA	1516	
QY	1231	TCGTGGGTGGGCAAGAGCACTAGCAGGGGAAAGTATGATTATTTACCGCTGGGGGAA	1290	
DB	1515	GGTGGGTCTCTACTCTGAAGCGCTAGCTGGTGAAGGCTTAATTCGTAACGGCTTGGTGGTA	1456	
QY	1291	TCGATTTACACACCCACTTTCTTCTCCACAAACAAATTCCTACCGCTCTAGCCCAATGGCG	1350	
DB	1455	TTGACACACATCCACTTCATTTTACCCCCCAACAAATCCCTACAGCTTTTGAAGCGGTG	1396	
QY	1351	TTTAAACCAATGTTTGGAGCGGCAAGGTCCTGTGATGGCAAGATGCGACTACTATATCA	1410	
DB	1395	TAAACCACTGATTTGGTGGGAAACCGGTCCTGCTGATGGCACTTAATTCGACTACTATCA	1336	
QY	1411	CTCCGGGCAAAATGGAACCTTGCACCGCATGTTGCGCCGACGAGAGTATTTCTATGAATG	1470	
DB	1335	CTCCAGGTAGAAGAACTTAAATTTGATGCTCAGAGCGGCTGAAGAAATATTTCTATGAAC	1276	
QY	1471	TGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAAAGTAGAAG	1530	
DB	1275	TAGGTTTCTTAGCTAAAGGTACGCTTTTAAACGCGAGCTTAGCGGATCAAAATGAAG	1216	
QY	1531	CGGGCGCATTTGGTTTAAATTTGCATGAAGACTTGGGGCAACAAACCAAGTGCATCGATC	1590	
DB	1215	CGGTGCGATTTGGCTTTAAAAATCCACGAAGACTTGGGGCACCACTCTCTTCGCAATCAATC	1156	
QY	1591	ACTGCTTGAGCTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGATACAGTCA	1650	
DB	1155	ATGCGTTAGACGTTGCGGCAAAATACGATGTGCAAGTCGCTATTCACACAGACACTTTGA	1096	
QY	1651	ATGAGCAGGTTATGTAGATGACCCCTTAAATGCAATGAACGGCGCGCCATCCATGCCT	1710	
DB	1095	ATGAAGCCGTTGCTGTGGAAGACACTATGGCAGCTATTGCGGACGCACTATGCACACTT	1036	
QY	1711	ACCACATTGAGGAGCGGGTGGAGGACACTCACTGATGTTATCAACATGGCAGGCGAGC	1770	
DB	1035	TCCACACTGAAGCGCTGGTGGCGGACACGCTCCTGACATTTATTAAGTGGCGGTGAAC	976	
QY	1771	TCAATATTTACCTCTCCACACCCCACTATTCCCTATACCATTTAATACGGTTTCAG	1830	
DB	975	ACAACATCTCTACCCGTTTCCACTTAACCCCACTATCCCTTTTCACTGTGTAATACAGAACCG	916	
QY	1831	AACACTTTAGACATGCTCATGACATGCAACCACTAGACAAACGATCCGCGAGGATTTAC	1890	
DB	915	AACATATGACATGCTCATGTTGTGCCAACACTTTGGATAAAGCAATTAAGAGATGTCC	856	
QY	1891	AATTTTCTCAAGCCCGTATCCGCCCGGCTCTATCCGGCTGGAAGATGTCTCCATGATA	1950	
DB	855	AGTTTCGCTGATTCAAGGATCCGCCCTCAAAACCATTTGCGGCTGAAGACACTTTTGCATGACA	796	
QY	1951	TGGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGCAAGTGA	2010	
DB	795	TGGGGATTTTCTCAATCAACCCAGTTCTGACTCTCAAGCGATGGGTGCTGTGGGTGAAGTTA	736	
QY	2011	TTCTCTCGAATTTGGCAGACTGCGGATAAGAAATAAAAAGAAATTTGGTAAAGCTTCTCTGAAG	2070	
DB	735	TCATAGGACTTTGGCAAAACAGCTGACAAAAACAAAAAAGAAATTTTGGCGGCTTTGAAGAG	676	
QY	2071	ATGGCAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACC	2130	
DB	675	AAAAAGCGGATACGACAACTTCAGGATCAACCGTACTTGTCTTAATATACCACTTAACC	616	
QY	2131	CCGCTTTGACCCACCGCGTGAGCGATATATTCGGCTCTGTGGAAAGAGGGGCAAGATTCGCCG	2190	

Db	615	CACGATCGCTCATGGGATTAGCGGATATAGCGTTCTGTAGAGTGGCGCAAGTGGCTG	556	FEATURES
Qy	2191	ACTTGGTGTGGTGAATCTCGCTCTTTTGGCGCTAAACCCAAATCGTATCAAAAGCGC	2250	source
Db	555	ACTTGGTATTGTGAGTCCAGCATCTTTGGCGTGAACCCACATCATCAAGGCG	496	
Qy	2251	GTATGGTGTCTTCTCTGAATATGGGCGATTCTAACGGTCTGTGCCACTCCCAACCGG	2310	gene
Db	495	GATTCAATTCGTTTAAGCCAAATGGGCGATGCGAAGCGTTCTATCCCTACCCCAACCGG	436	CDS
Qy	2311	TTTATTACCGCGAATGTTTGGGCATCAGCGAAGCGGAATTTGACACGATCATCTT	2370	
Db	435	TTTATTACAGAGAAATGTTGCTCACCATGGTAAAGCTAAATACGATGCAACATCACTT	376	
Qy	2371	TTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAG	2430	
Db	375	TTGTGTTCTCAGCGGCTTATGACNAAGGCATTAAAGAGNATTAGGGCTTGAAGACAG	316	
Qy	2431	TTCTACCGGTCAAAATCTCCGTPAACATCACCAAGAAAGCTTCAAGTTCAACGACAAA	2490	
Db	315	TGTTGCGGCTTAAATTCGAGAAACATCACTAAAAAGACATGCAATTCATGACACTA	256	
Qy	2491	CGCAAAATACCGTCGATCCGAAACCTTCGAGTCTTTGTAGATGCGAACTCTGCA	2550	
Db	255	CGCTCACATGAAGTCAATCTCGAAACTTACCATGTGTTCTGGATGCGCAAGATCA	196	
Qy	2551	CTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTAGG	2604	
Db	195	CTTCTAAACAGCTAATAAGTAGCTTGCTGCTCACTCTTACATTTTCTAGG	142	
RESULT 14				
AE000529/c				
LOCUS	AE000529	10903 bp	DNA	linear
DEFINITION	Helicobacter pylori 26695 section 7 of 134 of the complete genome.			
ACCESSION	AE000529	AB000511		
VERSION	AE000529.1	GI:2313152		
KEYWORDS				
SOURCE	Helicobacter pylori 26695			
ORGANISM	Helicobacter pylori 26695			
REFERENCE				
AUTHORS	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.			
	1 (bases 1 to 10903)			
	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.			
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori			
JOURNAL	Nature 388 (6642), 539-547 (1997)			
PUBMED	9252185			
REFERENCE				
AUTHORS	2 (bases 1 to 10903)			
	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Wathey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.			
	Direct Submission			
TITLE	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL				
PUBMED				
REFERENCE				
AUTHORS	3 (bases 1 to 10903)			
	White, O.			
TITLE	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL				

Location/Qualifiers

1. 10903
/organism="Helicobacter pylori 26695"
/mol_type="genomic DNA"
/strain="26695"
/db_xref="taxon:85962"
complement (127. .1836)
/gene="HP0072"
complement (127. .1836)
/gene="HP0072"
/note="similar to GB:M60398 GB:X17079 SP:P14917 PID:149011 GB:AE000511 percent identity: 100.00; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="urease beta subunit (urea amidohydrolase) (ureB)"
/protein_id="AAD07143.1"
/db_xref="GI:2313153"
/translation="MKKISRKEYSMYGPPTTGDKVRGLDGLIAEVEHDYTIYBELK FGCKTLREGMSQSNPSKEELDIITNALIVDTYIKADIGIKDGKGIAGKGGNK GMDGVKNLSVGPATEALAGEGLIVTAGGIDTHIFISPOQIPTAFASGVTWIGGG TGPADGTNATITPGRNLKWLRAABEYSMNLGFLAKGNASNDASLADQIEAGAIGF KIHEDWGTTPSAINHALDVADKYDQVAIHDTILNEAGCVDNWAALAGRTMHTFHE GAGGHPADIIKVAGSHNLPASTNPTIPTVNTAEHMDMLMVCCHLADKSKEDVQF ADSRIRPQTAAEDTLHDMGIFSITSSDSQAMGRVGEVITRTWQADKNKKEFGRLKE EKGNDNFRIKRYLSKYTINPAIAGHISVEYGVKADLWSPAFVGVKPNMI KGSFIALSQMGDANASIPTPQPVYVREMFAGHKAKYDANITFVSOAYDKGIKEELG LERQVDPVKNCRNITKKDMQNDTTAHIEVNPETYHYVVDGKEVTSKPANKVSLQLF SIF"
complement (1840. .2556)
/gene="HP0073"
complement (1840. .2556)
/gene="HP0073"
/note="similar to SP:P14916 PID:149010 PID:43634 GB:AE000511 PID:1567116 percent identity: 100.00; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="urease alpha subunit (urea) (urea amidohydrolase)"
/protein_id="AAD07144.1"
/db_xref="GI:2313154"
/translation="MKLTPEKLDKMLHVAGELAKKKEKGIKLNYYEAVALISAHIM EBARAKKTAABELMQEGRLLKPDVMDGVASMIHEVGIEAMEPDGTLVHTHTPEIA NGKLVPEGLFKNEDITINEGKAVKVKVKNVGRDPRVQIGSHFFFEVNRCLDFDREK TKGRLDIAGSTAVREFPEGEESVELIDIGNRRIFGFNALVDQRADNESKKIALHRA KERGFHAKSDDDNYVTKIE"
complement (2809. .2881)
/gene="tRNA-Val-2"
complement (2809. .2881)
/gene="tRNA-Val-2"
/product="tRNA-Val"
/note="codon recognized: GUC"
complement (2902. .3375)
/gene="HP0074"
complement (2902. .3375)
/gene="HP0074"
/note="similar to GB:M60398 SP:P25178 PID:149009 GB:AE000511 PID:2313155 percent identity: 100.00; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="lipoprotein signal peptidase (lspA)"
/protein_id="AAD07145.1"
/db_xref="GI:2313155"
/translation="MLKTTKSLVFMGVFFLIFGVDDQAIKVAILEGFRVRESLMIDIV LVPNKGVASLSLFLGGLKYLIQLILLGLIFLMQRQRELFKHAISEFVGVGAGVSN VLDRFVHGGVVDVYIYHGFDFAFNFDVMDVGVGVLLKQFFFKQKQNKKA"
complement (3369. .4706)
/gene="HP0075"
complement (3369. .4706)
/gene="HP0075"


```
Db      1803  ATGTATGGCCCTACTACAGCGCGATAAAGTGAAGATTTGGGCGATACAGACTTTGATCGCTGAA 1744
Qy      987   GTAGAACATGACTATATACCACTATATGCGGAAGAACTTTAAATTTTGGCGGGTAAAACTATC 1046
Db      1743  GTAGAACATGACTATACCACTTTATGCGGAAGAGCTTTAAATTCGGTGGCGGTAAAAACCTTG 1684
Qy      1047  CGTGAAGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCACT 1106
Db      1683  AGAAGAGGATGAGCAATCAACACCCCTAGCAAGAGAAATTTGGATCTTAATCATCACT 1624
Qy      1107  AACGCGATGATATCACTACACCGGGATTTACAAAGCCGACATTCGGGATTTAAAAAGCGC 1166
Db      1623  AACGCTTTAATCGTGGATTACACCGGTATTTTATAAGCGGATATGGTATTAAGATGCG 1564
Qy      1167  AAAATCCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCAT 1226
Db      1563  AAAATCGCTGGCATTTGGTAAAGCGGTACAAAGACATGCAAGATGGCGTTAAAAACAAT 1504
Qy      1227  ATGGTCGTGGGTGTGGGCACAGAAGCACTAGCAAGGGAAAGGTATGATTTATACCGCTGGG 1286
Db      1503  CTTAGCGTAGGTCCTGCTACTGAAGCCTTTAGCGGTGAAGGTTTGATCGTAACGTGCTGT 1444
Qy      1287  GGAATCGATTACACACCACTCTCTTCTCCACAACAATTCCTACCGCTCTAGCCAT 1346
Db      1443  GGTATTGACACACATCCACTTCAATTTCAACCCCAACAAATCCCTACAGCTTTTGAAGC 1384
Qy      1347  GCGTTTACAACCATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGGCACTACT 1406
Db      1383  GGTGTAACAACCATGATTTGGTGGCAACTGGTCTGCTGATGGCACTAATGGCACTACT 1324
Qy      1407  ATCACTCCGGGCAAAATGGAATTTGACCGCATGTTGCGGCGAGCAGAGAGTATTTCTATG 1466
Db      1323  ATCACTCCAGGCAGAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGATATTTCTATG 1264
Qy      1467  AATGGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAAGTA 1526
Db      1263  AACTTAGGTTTCTTGGCTAAAGTAACTACGCTTTTAAACGCGAGCTTAGCCGATCAAT 1204
Qy      1527  GAAGCGGCGCATTTGTTTAAATTTGATGATGAGCTGGGCGACACACCAAGTCGATC 1586
Db      1203  GAAGCTGGTGGCATTTGGCTTTAAATTCACGAGACTTGGGCGACCACTCTTCTTGAATC 1144
Qy      1587  GATCACTGCTTGAGCGGTGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db      1143  AATCATCGCTTAGATGTTGCAGCAATACGATGTGCAAGTGCCTATCCACACAGACAT 1084
Qy      1647  GTCAATGAGCGAGGTTATGATGATGACACCCCTAAATGCAATGAACCGGCGCGCATCAT 1706
Db      1083  TTGAATGAAGCCGTTGCGTGGAGACACTATGGCAGCTATTGCGGCGCACTATGACAC 1024
Qy      1707  GCTTACCACATTTGAGGAGCGGTGGAGACACTCACTCATGATGTTATCACCATGGCAGCG 1766
Db      1023  ACTTTTCCACACTGAAGGTCTGGCGGCGACACGCTCTCTGATATTTAAAGTAGCTGT 964
Qy      1767  GAGCTCAATATTCTACCCCTCTCCACCCCGCCACTATTCCCTATACCAATTAATACGGTT 1826
Db      963  GAACACAACTTTCTCCCGCTTCCACTAACCCCACTATCCCTTCACTGTGAATACAGAA 904
Qy      1827  GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGAT 1886
Db      903  GCAGAACACATGGACATGCTTATGGTGTGCCACCACTTGGATAAAGCATTTAAAGAGAT 844
Qy      1887  TTACAAATTTTCAAGCGGTATCCGCCCCGGCTCTATCGGGCTGAAGATGTCCTCAT 1946
Db      843  GTTCAGTTCGCTGATTTCAAGGATTCGCCCCCTCAAAACCAATTTGCGGCTGAAGACACT 784
Qy      1947  GATATGGGTGATCGCGATGACAGCTCGGATTTGCAAGCAATTCGCAAGGCGGTGCAGCGAA 2006
Db      783  GACATGGGGATTTTCTCAATACCAAGCTCTGACTCTCAAGCTATGGGTCTGTGGGTGAA 724
Qy      2007  GTGATTTCTCGAACTTGGCAGACTCGGGATAGAATAAAAAAGAAATTTGGTAAAGCTTCT 2066
Db      723  GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAAAAAGAAATTTGCGCGCTTGAA 664
```

```
Qy      2067  GAAGATGGCAAGATAAACGATAATTTCCGATTAAAGCGCTACATCTCAAATAACATATC 2126
Db      663  GAAGAAAAAGCGGATAAACGAACTTTAGGATCAACGCTACTTGTCTAAATAACACCAT 604
Qy      2127  AACCCCGCTTTGACCCACGCGGTGAGGAGTATATCGGCTCTGTGGAGAGGGCAAGATC 2186
Db      603  AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTCTGTAGAGTGGGCAAGTG 544
Qy      2187  GCCGACTTGTGTGTGGAAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAA 2246
Db      543  GCTGACTTGTATTTGGAGTCCGCACTTTTGGCGTAAACCCCAACATGATCATCAAA 484
Qy      2247  GGCGGTATGGTGTCTTCTGAAATGGGCGATTTCTAACCGCTGTGTGCCACATCTCCCAA 2306
Db      483  GGCGGTTCATTCGCTGAGTCAATGGGTGACGCAACGCTTCTATCCCTACCCCAAA 424
Qy      2307  CCGGTTTATTTACCGCAAAATGTTTGGGCATCAGCGCAGCGGAAATTTGACACAGATC 2366
Db      423  CCAGTTTATTTACAGAGAAATGTTTCGCTCATCATGGTAAAGCCAAATACGATGCAACATC 364
Qy      2367  ACTTTTGTTCCAAAGTCGCTATGAAATGGGTGAAAGAAAGCTGGGCTTAGAGGCG 2426
Db      363  ACTTTTGTGCTCAGCGCTTATGACAAAGGATTTAAAGAGATTTAGGCTTTGAAGA 304
Qy      2427  CRAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db      303  CRAAGTGTTCGCGTAAAAAATTCGAGAAACATCACTAAAAAAGACATGCAATTTCAACGAC 244
Qy      2487  AAAACGGCAAAATCAACGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
Db      243  ACTACGCTCACAATTGAAGTCAATCTGAAACTTACCACTGTGTCTGTGGATGGCAAGAA 184
Qy      2547  TGCACCTCTAAACCCACCTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db      183  GTAACTTCTAAACACGCAATAAAGTGAGCTTGGCGCACTCTTTAGCAITTTCTAGG 126
```

RESULT 15
AR054309

```
LOCUS      AR054309      2735 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5837240.
ACCESSION  AR054309
VERSION     AR054309.1 GI:5979886
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 2735)
AUTHORS     Lee, C.K., Monath, T.P., Ackerman, S.K., Thomas, W.D., Soman, G.,
            Kleanthous, H., Weltzin, R.A., Pappo, J., Ermak, T., Guirakho, P.,
            Bhagat, H. and Sussman, I.
TITLE       Multimeric, recombinant urease vaccine
JOURNAL     Patent: US 5837240-A 1 17-NOV-1998;
FEATURES    Location/Qualifiers
            source
            1..2735
            /organism="unknown"
            /mol_type="unassigned DNA"
```

ORIGIN

```
Query Match      37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred. No. 2.4e-204;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

Qy      182  AAAATTTAAACACAGGAGTAATAGGTGAAACTCACAACCCCAAGAGCAAGAAAGTTCTTG 241
Db      78  ACACCTTTAAGATAGGAGATGAGATGAAACTCACCCCAAAAGAGTTAGATAGTGTATG 137
Qy      242  TTATATTTGCGGGCGAAGTGGCTAGAACGCAAGCGCAAGCAGAGGGCTTTAAGCTCAACCA 301
Db      138  CTCACACTCGCTGGAGATTTGGCTAAAAAACGCAAGAAAGAGGCATTTAAGCTTAACCT 197
Qy      302  CCGAAGCCATTGCTTTACATTAGTGCCCATATTATGCAAGCGCGCGGTGGAAGAAAAA 361
```


Qy	2487	AAAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC	2546
Db	2415	ACTACTGCTCACATTGAAGTCAATCCTGAAACTTTACCATGTGTTTCGTGGATGGCAAGAA	2474
Qy	2547	TGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCCAGCGCTACACTTTCTTCTAGG	2604
Db	2475	GTAACCTTCTAAACCCAGCCAAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG	2532

Search completed: November 29, 2005, 04:30:37
 Job time : 14066 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:56:44 ; Search time 1565 Seconds
(without alignments)
12277.514 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 tggtagatttccarctt.....aaaaagtagagccacag 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query		DB	ID	Description
	Match	Length	Match	Length			
1	2880.6	99.9	2883	7	ADJ58237	Adj58237 Urease su	
2	2228	77.3	2452	7	ADJ58249	Adj58249 UreaseXY	
3	2158.6	74.9	2405	7	ADJ58240	Adj58240 UreaseXY	
4	2158.6	74.9	2407	7	ADJ58246	Adj58246 UreaseXY	
5	1948	67.6	2183	7	ADJ58243	Adj58243 UreaseXY	
6	1184.4	41.1	8407	12	ADQ37847	Adq37847 H. bizzoz	
7	1075.2	37.3	2735	2	AAT44351	Aat44351 H. pylori	
8	1062.6	36.9	2619	2	AAQ90180	Aaq90180 Helicobac	
9	1062.6	36.9	2619	2	AAT45680	Aat45680 H. felis	
10	1061	36.8	2619	2	AAQ75319	Aaq75319 Urease ur	
11	1058.6	36.7	4824	5	AEBS5130	Aeb55130 Salmonell	
12	1058.6	36.7	4824	5	AEBS5112	Aeb55112 Salmonell	
13	1053.6	36.5	2767	2	AAQ12485	Aaq12485 DNA encod	
14	940.8	32.6	1710	12	ADQ37850	Adq37850 H. bizzoz	
15	902.4	31.3	1719	8	ABA00816	Abao0816 H. felis	
16	885.8	30.7	1815	2	AAx14230	Aax14230 H. pylori	
17	884.8	30.7	1710	8	ACA34557	Aca34557 Prokaryot	
18	881.8	30.6	1707	13	ADU05358	Adu05358 DNA encod	
19	881.8	30.6	1707	13	ADU05359	Adu05359 DNA encod	

20	880	30.5	1710	9	AAL60579	Aal60579 Helicobac
21	878.4	30.5	1710	2	AAQ04329	Aaq04329 Probe for
22	876.8	30.4	1717	8	ABA00817	Abao0817 H. pylori
23	855.6	29.7	2385	2	AAV62460	Aav62460 Helicobac
24	734.6	25.5	2341	8	ACA45080	Aca45080 Prokaryot
25	693.4	24.1	2400	2	AAV19002	Aav19002 Klebsiell
26	693.4	24.1	4768	3	AAZ92097	Aaz92097 Klebsiell
27	690.6	24.0	2439	8	ACA46520	Aca46520 Prokaryot
28	664.4	23.0	110000	2	AAT42063	Continuation (6 of
29	658.6	22.8	1698	8	ACA45232	ACA45232 Prokaryot
30	653	22.7	1710	10	ADF01572	Adf01572 Bacterial
31	634.8	22.0	3919	13	ADT05424	Adt05424 Haemophil
32	634.8	22.0	85814	13	ADT05644	Adt05644 Haemophil
33	630.6	21.9	8729	9	ACD19244	Act19244 E. coli 0
34	630.6	21.9	86248	10	ADC00087	Adc00087 Enterohae
35	630.6	21.9	87563	9	ACD19044	Act19044 E. coli 0
36	627.6	21.8	5966	2	AAV74333	Aav74333 Staphyloc
37	625.4	21.7	1719	8	ACA34156	Aca34156 Prokaryot
38	617.8	21.4	1878	11	ACH96250	Ach96250 Klebsiell
39	612	21.4	1701	8	ACA35953	Aca35953 Prokaryot
40	612	21.2	3164	4	AAH54464	Aah54464 S. epider
41	612	21.2	3234	4	AAH54489	Aah54489 S. epider
42	609.6	21.1	1704	8	ACA25712	Aca25712 Prokaryot
43	603.4	20.9	1704	8	ACA23584	Aca23584 Prokaryot
44	602.8	20.9	1719	8	ACA53973	Aca53973 Prokaryot
45	591.4	20.5	1698	8	ACA20910	Aca20910 Prokaryot

ALIGNMENTS

RESULT 1
ADJ58237
ID ADJ58237 standard; DNA; 2883 BP.
XX AC ADJ58237;
XX DT 06-MAY-2004 (first entry)
DE Urease subunit polypeptide complex encoding sequence.
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX OS Helicobacter felis.
XX FH Key Location/Qualifiers
CDS 206..886
FT /*tag= a
FT /*product= "urease polypeptide complex"
CDS 897..2603
FT /*tag= b
FT /*product= "urease polypeptide complex"

XX EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX DR WPI; 2002-124384/17.
XX DR P-PSDB; ADJ58238, ADJ58239.
XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX PS Claim 1; SEQ ID NO 1; 76pp; English.

XX	The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an urease subunit polypeptide complex encoding sequence of the invention.									
SQ	Sequence	2883 BP;	871 A;	642 C;	666 G;	698 T;	0 U;	6 Other;		
Query Match 99.9%; Score 2880.6; DB 7; Length 2883;										
Best Local Similarity 100.0%; Pred. No. 0;										
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
QY	1	RGRAGATTTCCARCACCTCAAGCACATATTCATCTGTGTGGTGCTAAATTCR	60							
DB	1	RGRAGATTTCCARCACCTCAAGCACATATTCATCTGTGTGGTGCTAAATTCR	60							
QY	61	ACTTGTAAATRCATTAATTAATTTTAAATAATTAATTAATTAATTAATTA	120							
DB	61	ACTTGTAAATRCATTAATTAATTTTAAATAATTAATTAATTAATTAATTA	120							
QY	121	TTACTTATATAAAGGTTAAATAAGTTAACGAAATTAGGACTATAATCCCATTCGCTT	180							
DB	121	TTACTTATATAAAGGTTAAATAAGTTAACGAAATTTAGGACTATAATCCCATTCGCTT	180							
QY	181	TAAATTTTAAACAAGAGTAAATAGGTGAACCTCACACCAAGAGCAAGAAAGTTCTT	240							
DB	181	TAAATTTTAAACAAGAGTAAATAGGTGAACCTCACACCAAGAGCAAGAAAGTTCTT	240							
QY	241	GTTTATATTATCGGGGGAAGTGTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCA	300							
DB	241	GTTTATATTATCGGGGGAAGTGTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCA	300							
QY	301	ACCCGAAGCCATTGCTTACATTAATTAATTAATTAATTAATTAATTAATTA	360							
DB	301	ACCCGAAGCCATTGCTTACATTAATTAATTAATTAATTAATTAATTAATTA	360							
QY	361	AACCGTGTCCAGCTTATGAAGAGTGCATGCACTTTTGAAAGAGATGAATGTC	420							
DB	361	AACCGTGTCCAGCTTATGAAGAGTGCATGCACTTTTGAAAGAGATGAATGTC	420							
QY	421	CGGGTGGGTAATATGTTTCCCATCTAGTGTAGAGCCACCTTCTCATGTGTACGA	480							
DB	421	CGGGTGGGTAATATGTTTCCCATCTAGTGTAGAGCCACCTTCTCATGTGTACGA	480							
QY	481	ACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGCGGAAGTGA	540							
DB	481	ACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGCGGAAGTGA	540							
QY	541	ATTTGGTGGATTAAGACATCGAGCTCAATGCGGCAAGAGTAACCGAATGAGGT	600							
DB	541	ATTTGGTGGATTAAGACATCGAGCTCAATGCGGCAAGAGTAACCGAATGAGGT	600							
QY	601	TACTTAATGAAGGCTTAATTCCTTTCATGAGTGGTGGTGGTGGTGGTGGTGGT	660							
DB	601	TACTTAATGAAGGCTTAATTCCTTTCATGAGTGGTGGTGGTGGTGGTGGTGGT	660							
QY	661	CAAGGCACTAAATTCGATCGTGAANAAGCCTATGSCAAGCCCTAGATATTCCTCTGG	720							
DB	661	CAAGGCACTAAATTCGATCGTGAANAAGCCTATGSCAAGCCCTAGATATTCCTCTGG	720							
QY	721	CAACAGCGTACGANTTGGGGCAGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780							
DB	721	CAACAGCGTACGANTTGGGGCAGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780							
QY	781	CAGTAAATAAGTGAATGCGCATGAACGGCTTGTGAATAACATCCGGATGAACGCCATAA	840							
DB	781	CAGTAAATAAGTGAATGCGCATGAACGGCTTGTGAATAACATCCGGATGAACGCCATAA	840							
QY	841	ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTTATCAAGTAAGAGAGACTCCCATGA	900							

DB	841	ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTTCAAGTAAGAGAGACTCCCATGA	900							
QY	901	AAATGAAAAAACAAGATAATGTAAATACCTACGAGCCACCACAAAGGGGATAAAGTCGCT	960							
DB	901	AAATGAAAAAACAAGATAATGTAAATACCTACGAGCCACCACAAAGGGGATAAAGTCGCT	960							
QY	961	TAGAGATACCGATCTTTTGGCGAGAACTAGAACATGACTATACCACTATCGGGAAGAAC	1020							
DB	961	TAGAGATACCGATCTTTTGGCGAGAACTAGAACATGACTATACCACTATCGGGAAGAAC	1020							
QY	1021	TTAAATTTGGCGGGTAAACATCTCGTGAGGGTATGGGTGAGGAGTATAGCCCTGATG	1080							
DB	1021	TTAAATTTGGCGGGTAAACATCTCGTGAGGGTATGGGTGAGGAGTATAGCCCTGATG	1080							
QY	1081	AAAAACCCCTAGATTTTATCATCACTAACCGCGATGATTCGACTACACCCGGGATTTACA	1140							
DB	1081	AAAAACCCCTAGATTTTATCATCACTAACCGCGATGATTCGACTACACCCGGGATTTACA	1140							
QY	1141	AAGCGCATTTGGGATTTAAAAACGGCAAAATTCATGGCATTTGGCAAGCAGGAAACAAG	1200							
DB	1141	AAGCGCATTTGGGATTTAAAAACGGCAAAATTCATGGCATTTGGCAAGCAGGAAACAAG	1200							
QY	1201	ACATGCAAGATGGCGTAAGCCCTCATATGTCGTCGGTGTGGGCACAGACACTAGCAG	1260							
DB	1201	ACATGCAAGATGGCGTAAGCCCTCATATGTCGTCGGTGTGGGCACAGACACTAGCAG	1260							
QY	1261	GGGAAGGTATGATTTTACCGCTGGGGAATTCGATTCACACACCCACCTTCCTTCTCCAC	1320							
DB	1261	GGGAAGGTATGATTTTACCGCTGGGGAATTCGATTCACACACCCACCTTCCTTCTCCAC	1320							
QY	1321	AACAATTTCCCTACCGCTCTAGCCAAATGGCGTTAACAACATGTTTGGAGCGGCAACGTC	1380							
DB	1321	AACAATTTCCCTACCGCTCTAGCCAAATGGCGTTAACAACATGTTTGGAGCGGCAACGTC	1380							
QY	1381	CTGTAGATGGCAAGATGGCACTACTATCACTCGGGCAATGGAACCTTGACCCGATGT	1440							
DB	1381	CTGTAGATGGCAAGATGGCACTACTATCACTCGGGCAATGGAACCTTGACCCGATGT	1440							
QY	1441	TGCGCGCAGCAGAGAGTATTTCTATGATGCGGCTTTTGGGCAAAAGGCAATAGCTCTA	1500							
DB	1441	TGCGCGCAGCAGAGAGTATTTCTATGATGCGGCTTTTGGGCAAAAGGCAATAGCTCTA	1500							
QY	1501	GCAAAAAACAATTTGTAGAACAGTAGAAGCGGCGGATTTGGTTTTAAATTTGCATGAAG	1560							
DB	1501	GCAAAAAACAATTTGTAGAACAGTAGAAGCGGCGGATTTGGTTTTAAATTTGCATGAAG	1560							
QY	1561	ACTGGGGCAACACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATAGCATG	1620							
DB	1561	ACTGGGGCAACACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATAGCATG	1620							
QY	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCTAA	1680							
DB	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCTAA	1680							
QY	1681	ATGCAATGAACCGGCGGCCATCCATGCTTACCACTTGGAGGAGCGGTTGGAGGACACT	1740							
DB	1681	ATGCAATGAACCGGCGGCCATCCATGCTTACCACTTGGAGGAGCGGTTGGAGGACACT	1740							
QY	1741	CACCTGATTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA	1800							
DB	1741	CACCTGATTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA	1800							
QY	1801	CTATTTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGATGCCACC	1860							
DB	1801	CTATTTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGATGCCACC	1860							
QY	1861	ACCTAGACAAACCGATCCGCGAGGATTTTCAATTTTCTCAAGCCGATATCCGCGCGCT	1920							
DB	1861	ACCTAGACAAACCGATCCGCGAGGATTTTCAATTTTCTCAAGCCGATATCCGCGCGCT	1920							
QY	1921	CTATCGCGCTGGAAGATGCTTCCATGATATGGGTGATCGCGATGAACAGCTCGGATTT	1980							

Db 1921 CTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATT 1980
Qy 1981 CGCAAGCAATGGCGCTGCAGGCGAAGTGAATCTCTCGAACTTGGCAGACTGCGGATAAGA 2040
Db 1981 CGCAAGCAATGGCGCTGCAGGCGAAGTGAATCTCTCGAACTTGGCAGACTGCGGATAAGA 2040
Qy 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
Db 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
Qy 2101 AGCGGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
Db 2101 AGCGGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
Qy 2161 TCGGCTCTGTGAAGAGGCAAGATCGCCGACTTGGTGTGTGGAATCTGCGCTTTTGG 2220
Db 2161 TCGGCTCTGTGAAGAGGCAAGATCGCCGACTTGGTGTGTGGAATCTGCGCTTTTGG 2220
Qy 2221 GCGTAAACCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGATT 2280
Db 2221 GCGTAAACCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGATT 2280
Qy 2281 CTAAAGCGCTGTGCGCACTCCCAACCGGTTTATTACCGCAATGTTTGGGCATCAG 2340
Db 2281 CTAAAGCGCTGTGCGCACTCCCAACCGGTTTATTACCGCAATGTTTGGGCATCAG 2340
Qy 2341 GCAAGCGGAAATTTGACACGAGCATCACTTTTGTTCCTCAAGTCGCTATGAAATGGCG 2400
Db 2341 GCAAGCGGAAATTTGACACGAGCATCACTTTTGTTCCTCAAGTCGCTATGAAATGGCG 2400
Qy 2401 TGAAGAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACTGCCGTAAACATCA 2460
Db 2401 TGAAGAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACTGCCGTAAACATCA 2460
Qy 2461 CCAAGAAAGCTTCAAGTTCAACGACAAACGCGCAAAATCAACCGTCGATCGCAAAACCT 2520
Db 2461 CCAAGAAAGCTTCAAGTTCAACGACAAACGCGCAAAATCAACCGTCGATCGCAAAACCT 2520
Qy 2521 TCGAGGTCTTTAGATGCGCAACTCTGACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Db 2521 TCGAGGTCTTTAGATGCGCAACTCTGACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Qy 2581 CCCAGCGCTACACTTTCTTTAGGCAATGCCCCCTTTGGGGCGCAGGTATTTTAGGAA 2640
Db 2581 CCCAGCGCTACACTTTCTTTAGGCAATGCCCCCTTTGGGGCGCAGGTATTTTAGGAA 2640
Qy 2641 TCTTCATCAACGCACTCGCAATCGGTCTTGGGTGTGCGATCGTGTGCGATCGTGTGCGAT 2700
Db 2641 TCTTCATCAACGCACTCGCAATCGGTCTTGGGTGTGCGATCGTGTGCGATCGTGTGCGAT 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCAATTTTAAATTAATTTAATTTCTTATTAATTAATTAAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCAATTTTAAATTAATTTAATTTCTTATTAATTAATTAAT 2760
Qy 2761 TATGCCCTCTATTTTAAAGGAGAAATATGCGTAGTCTTTTGGTATTTGCTATGTTGGGG 2820
Db 2761 TATGCCCTCTATTTTAAAGGAGAAATATGCGTAGTCTTTTGGTATTTGCTATGTTGGGG 2820
Qy 2821 TTGTTTGGTGTGGGCGCAAGGGTATTGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGGGCGCAAGGGTATTGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 2
ADJ58249
ID ADJ58249 standard; DNA; 2452 BP.
XX
AC ADJ58249;
XX

DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #4.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
CDS 48..728
FT /*tag= a
FT /product= "urease protein"
FT 739..2445
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PP 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
XX WPI; 2002-124384/17.
DR P-PSDB; ADJ58250, ADJ58251.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 13; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;
Query Match 77.3%; Score 2228; DB 7; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGACTATATCCCATTCCTTTAAATTTTAAACAAAGGAGTATAGGTGAACTCACAC 218
Db 1 AGACTATATCCCATTCCTTTAAATTTTAAACAAAGGAGTATAGGTGAACTCACAC 60
Qy 219 CCNAGAGCAAGAAAGTCTTCTTATTTATTCGGGCGAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCNAGAGCAAGAAAGTCTTCTTATTTATTCGGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGTGCCCATATTATGG 180
Qy 339 ACGAAGCGCGCTGGGAAACCCGTTCCCGAGCTTATGGAAAGTGCATGACATTTT 398
Db 181 ACGAGCGCGCTGCTGGCAAAAAACCGTTCCGGAACCTTATGGAAGAGTGTATGCATTTT 240
Qy 399 TGAANAAGATGAGTAATATCCCGGGTGGTAAATGTTTCCGATCTAGGTCTAGAAG 458
Db 241 TGAANAAGATGAGTAATATCCCGGGTGGTAAATATGTTCCCTGATTTGGGCGTGAAG 300
Qy 459 CCACCTTTCTGTGATGATCGAAACTTGTAACTGTGAATTGGCCCATCGAACCATGAGC 518

[illegible]

```
ID ADJ58240 standard; DNA; 2405 BP.
XX AC ADJ58240;
XX XX
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE UreaseX subunit encoding sequence #1.
XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX OS Helicobacter felis.
XX FH
XX FT Location/Qualifiers
XX FT 1. .681
XX FT /*tag= a
XX FT /product= "urease protein"
XX FT 692. .2398
XX FT /*tag= b
XX FT /product= "urease protein"
XX XX
XX PN EP1176192-A2.
XX XX
XX PD 30-JAN-2002.
XX PF
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR
XX PR 17-JUL-2000; 2000EP-00202565.
XX XX
XX PA (ALKU ) AKZO NOBEL NV.
XX PI
XX PI Kusters JG, Cattoli G;
XX XX
XX DR WPI: 2002-124384/17.
XX DR P-PSDB; ADJ58241, ADJ58242.
XX XX
XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX PT the diagnosis of Helicobacter felis infections and in the preparation of
XX PT vaccines.
XX PS Disclosure; SEQ ID NO 4; 76pp; English.
XX CC
XX CC The present invention relates to a novel Helicobacter felis urease X and
XX CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX CC used in the manufacture of vaccines against Helicobacter felis infections
XX CC and in diagnostic tests to detect antibodies against Helicobacter felis.
XX CC Helicobacter felis is difficult to grow so it is more convenient to use
XX CC the expression products of the genes encoding the urease X and Y subunits
XX CC in the manufacture of vaccines. The present sequence represents an
XX CC ureaseX subunit of the invention.
XX XX
XX SQ Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;
XX XX
XX Query Match 74.9%; Score 2158.6; DB 7; Length 2405;
XX Best Local Similarity 93.6%; Pred. No. 0;
XX Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
XX
XX 206 GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 265
XX 1 GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 60
XX
XX 266 AGAAGCGCAAGCAGAGGCGCTTAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT 325
XX
XX 61 AGAAGCGCAAGCAGAGGCGCTTAAGCTCAATCAACCCGAGCCATTGCTTACATTAGT 120
XX
XX 326 GCCCATATTATGGAAGAGCGCGCTTGAAAAAACCCTTGCCGAGCTTATGGAAGAG 385
XX 121 GCCCATATTATGGAAGAGCGCGCTTGAAAAAACCCTTGCTGAACTTATGGAAGAG 180
XX
XX 386 TGCATGCACTTTTGAAGAAAGATGAAGTAATGCCGGGGTGGGTAAATATGTTCCCGAT 445
XX 181 TGTATGCACTTTTGAAGAAAGATGAGGTGATGCCCGGTGCGGGAATATGGTCCCTGAT 240
XX
XX 446 CTAGGTGTAGAGCCACCTTTCTCTGATGTCGAAACTTGTAACTGTGAACTTGGCCCATC 505
```

```

XX 241 TTGGGCGTAGAGCCACTTTCCCGGATGGCACCACAACTCGTAACCGTGAATTGGCCCAT 300
XX
XX 506 GAACCCAGATGAGCACTTCAAAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
XX
XX 301 GAACTGTATGAACACACTTTAAAGCGGTAAGTGAATTTGGCTGTGATAAAGACATTTGAG 360
XX
XX 566 CTCAATGTCAGGCAAGAAAGTAACCGAACTTTGAGGTTACTAATGAAGGCGCTTAAATCTCTTG 625
XX
XX 361 CTCAACGCGGGTAAGGAAGTTACCGAGCTTTGAAGTTACCAACGAAGGACCTTAAATCTCTTG 420
XX
XX 626 CATGTGGGTAGCCATTTCCTCTTTGAAAGCTTAAACGAAGGCACTTAAATTCGATCGTGAA 685
XX
XX 421 CATGTGGGTAGCCATTTCCTCTTTGAAACCAACCAAGGCATTGAAATTCGATCGGGAA 480
XX
XX 686 AAAGCCCTATGGCAAAACGCTTAGATATTTCCCTCTGGCAACACGCTACCGATTGGGGCAGGA 745
XX
XX 481 AAAGCCCTATGGCAAAACGCTTAGATATTTCCCTCTGGCAACACGCTACCGATTGGGGCAGGA 540
XX
XX 746 CAAACCCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTAAAAAAGTGAATTTGGCATGAAC 805
XX
XX 541 CAAACCCGTAAGTGCAGTTAATCCCTCTTTGGCGTAGTAAAAAAGTGAATTTGGCATGAAC 600
XX
XX 806 GGCTTTGTGAATAACATCGCGGATGAACGCCATAAACAATAAAGCGCTTGCACAAAGCGGAAA 865
XX
XX 601 GGCTTTGTGAATAAATATTCGGACGAACGCGCATAAACACAAAGCACTAGACAAGGCAAAA 660
XX
XX 866 TCTCAGGATTATCAAGTAAGGAGACTCCCATGAAATGAAATAAACAAGTAATGTAAA 925
XX
XX 661 TCTCAGGATTATCAAGTAAGGAGACTCCCATGAAATGAAATAAACAAGTAATGTAAA 720
XX
XX 926 TACTACGGACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGCATCTTTGGGCGAGA 985
XX
XX 721 CACTACGGACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGCATCTTTGGGCGAGA 780
XX
XX 986 AGTAGAACATGACTATACCACTTATGCGCGAAGAACTTTAAATTTGGCCCGGTAAAACTAT 1045
XX
XX 781 AGTAGAACATGACTATACCACTTATGCGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT 840
XX
XX 1046 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTCATCAC 1105
XX
XX 841 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTCATCAC 900
XX
XX 1106 TAACGCGATGATTATGACTACACCCGATTTTCAAGCCGACATTCGGATTAAAAACGG 1165
XX
XX 901 CAAAGCGATGATTATGACTACACCCGATTTTAAAGCCGACATTTGTTATTTAAAAATGG 960
XX
XX 1166 CAAAATCCATGCGATTGGCAAGGCGAGGAAAACAAGGACATGCAAGATGCGGTAAAGCCCTCA 1225
XX
XX 961 CAAAATCCATGCGATTGGCAAGGCGAGGAAAACAAGGACATGCAAGATGCGGTAAAGCCCTCA 1020
XX
XX 1226 TATGTCGTGGGTGAGGCGACAGAACCACTAGCAGGGGAAAGGTATGATTTATCCGCTGG 1285
XX
XX 1021 TATGTCGTGGGTGAGGCGACAGAACCACTAGCAGGGGAAAGGTATGATTTATCCGCTGG 1080
XX
XX 1286 GGGATTCGATTACACACCCACTTCTCTTCTCCACACAAATTCCTACCCGCTCTAGCCAA 1345
XX
XX 1081 GGGATTCGATTACACACCCACTTCTCTTCTCCACAAATTCCTACCCGCTCTAGCCAA 1140
XX
XX 1346 TGGCGTTACAAACCATGTTGGAGGCGGACAGGTCTCTGTAGATGGCAAGTGGCACTAC 1405
XX
XX 1141 TGGTGTACAAACCATGTTGGAGGTGGCACAGGTCCGGTAGATGGCACGATGGCACAC 1200
XX
XX 1406 TATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGCGCAGCAGAGAGTATTTAT 1465
XX
XX 1201 CATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGCGCAGCTGAAGAGTATTTAT 1260
XX
XX 1466 GAATGCGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACCTTTAGAACAGT 1525
XX
XX 1261 GAATGCGGCTTTTGGGCAAAAGGCAATAGCTCTAGTAAAAAACAACCTTTAGAACAGT 1320
XX
XX 1526 AGAAGCGGCGGATTTGGTTTTTAAATTTGCAATGAAGCTGGGGGCAACACCAAGTCGAT 1585
```


Db 1321 AGAAGCGGCGGCGATTGGCTTTAAATTGCGATGAAGACTGGGCGACAAACACCAAGTCGGAT 1380
QY 1586 CGATCACTGCTTGGAGCGTGGCAGATGAATACAGATGTCGAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGGAGCGTGGCAGATGAATACAGATGTCGAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCTTAAATGCAATGAACCGGCGCGCCATCCA 1705
Db 1441 GGTCAATAGGCGAGGTTATGTAGATGACACCTTAAATGCGATGAACCGGCGCGCCATCCA 1500
QY 1706 TGCCTACCACTTAGGCGAGGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGG 1765
Db 1501 TGCCTACCACTTAGGCGAGGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGG 1560
QY 1766 CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT 1620
QY 1826 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA 1680
QY 1886 TTTACAAATTTCTCAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTCTCCA 1945
Db 1681 TCTCCAGTTTCTCAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTCTCCA 1740
QY 1946 TGATATGGGTGTGATCGCGATGACAACTCGGATTCGCAAGCAATGGGCGGTGCGAGCGA 2005
Db 1741 TGATATGGGTGTGATCGCGATGACAACTCGGATTCGCAAGCAATGGGCGGTGCGAGCGA 1800
QY 2006 AGTGATTTCTCGAACTTGGCAGACTCGGGAAGAATAAAGAAATTTGGTAAAGTTCC 2065
Db 1801 AGTGATTTCTCGAACTTGGCAGACTCGGGAAGAATAAAGAAATTTGGTAAAGTTCC 1860
QY 2066 TGAAGATGGCAAGATGAACATTAATTTCCGCATTAAGCGCTACATCTCCGAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATTAATGACAACTTCGCGATCAAAACGCTATATCTCCAAATACACCAT 1920
QY 2126 CAACCCCGCTTTCACCAACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db 1921 TAATCCCGCTTTCACCAACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGAAGAT 1980
QY 2186 CGCGGACTTGGTGTGGAATCTGCTCTTTTGGGTAAAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGGACTTGGTGTGGAATCTGCTCTTTTGGGTAAAAACCCCAAAATCGTGATCAA 2040
QY 2246 AGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACGCGTCTGTGCCCACTCCCA 2305
Db 2041 AGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACGCGTCTGTGCCCACTCCCA 2100
QY 2306 ACCGTTTATACCGGAAATGTTGGGATCAAGGCGCAAGGCGAAATTTGCACACCGCAT 2365
Db 2101 GCGGTTTATACCGGAAATGTTGGGATCAAGGCGCAAGGCGAAATTTGCACACCGCAT 2160
QY 2366 CACTTTTGTGTTCCAAAGTCGCTATGAAAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTGTTCCAAAGTCGCTATGAAAAATGGCGTGAAGAAAGAAAGCTAGGCTTAGAGCG 2220
QY 2426 CCAAGTTCTACCGGTCAAAAATCGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTGTACCGGTCAAAAATCGCGGTAAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
QY 2486 CAACACGGCAAAATCAACCGTCAATCCGAAACCTTCGAGGTCTTTGTAGATGCAAACT 2545
Db 2281 CAACACGGCGCATATCATGTGCGATCTTAAACCTTCGAGGTCTTTGTAGATGCAAACT 2340
QY 2546 CTGCACCTCTAAACCCACCTCGCAAGTCCCTCTAGCCAGCGCTACACTTCTTCTTAGGC 2605
Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCTCTAGCCCAACGCTACACTTCTTCTTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4

ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
PH Location/Qualifiers
FT CDS
FT /*tag= a
FT /product= "urease protein"
FT 693..2399
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PP 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR P-PSDB; ADJ58247, ADJ58248.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 10; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;
Query Match 74.9%; Score 2158.6; DB 7; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
Qy 206 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATGCGGGCGAAGTGGCT 265
Db 2 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATGCGGGCGAAGTGGCT 61
Qy 266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCATTTGCTTACATTAGT 325
Db 62 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCATTTGCTTACATTAGT 121
Qy 326 GCCCATATTATGGACGAGCGCGCGTGGAAAAAACCCGTTGCCAGCTTATGGAAGAG 385
Db 122 GCCCATATTATGGACGAGCGCGCGTGGAAAAAACCCGTTGCCAGCTTATGGAAGAG 181
Qy 386 TGCATGCACATTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGGTAATATGTTTCCCGAT 445

Db 1203 CATCACTCCGGGCAAAATGGAACTTGACCCGCATGTGTCGGCGCAGCTGAAGAGTATTCTAT 1262
QY 1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATCTGTAGAACAAAGT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATCTGTAGAACAAAGT 1322
QY 1526 AGAAGCGGCGGCGATTGGTTTAAATTCGATGAAGACTGGGGCACAACACCAAGTCCGAT 1585
Db 1323 AGAAGCGGCGGCGATTGGCTTTAAATTCGATGAAGACTGGGGCACAACACCAAGTCCGAT 1382
QY 1586 CGATCACTGCTTGGCGTGGCAGTGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1383 CGATCACTGCTTGGCGTGGCAGTGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1442
QY 1646 AGTCAATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCATCCA 1705
Db 1443 GGTCAATGAGCGAGGTTATGTGATGACACCCCTAAATGCAATGAACGGCGGCCATCCA 1502
QY 1706 TGCCTACCACTTGAAGGAGCGGCTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1765
Db 1503 TGCCTACCACTTGAAGGAGCGGCGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1562
QY 1766 CGAGCTCAATATCTACCTCTCCGACACCCCACTATTCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATATCTACCTCTCCGACACCCCACTATTCCTATACCATTAATACGGT 1622
QY 1826 TGCAGAACCTTAGACATGCTCATGATGCGACACCTAGACAAACGCGATCGCGAGGA 1885
Db 1623 TGCAGAACCTTAGACATGCTCATGATGCGACACCTAGACAAACGCGATCGCGAGGA 1682
QY 1886 TTTTACAATTTTCTCAAAGCGGTATCGGCCCGGCTCTATCGCGGCTGGAAGATGTGTCTCA 1945
Db 1683 TTTTACAATTTTCTCAAAGCGGTATCGGCCCGGATCTATTCGCGTGAAGATGTGTCTCA 1742
QY 1946 TGATATGGGTGTATCGGATGACAAAGCTCGGATTCGGAAGCAATGGGCGGTGACGGGA 2005
Db 1743 TGATATGGGTGTATCGGATGACAAAGCTCGGATTCGGAAGCAATGGGCGGTGACGGGA 1802
QY 2006 AGTGATTCCTCAACTTGGCAGCTGCGGATGAAGATGAAGATGAAGATTTGGTAAAGCTTC 2065
Db 1803 AGTGATTCCTCAACTTGGCAGCTGCGGATGAAGATGAAGATGAAGATTTGGTAAAGCTTC 1862
QY 2066 TGAAGATGCAAGATGAAGATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGTCAGATGAAGATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 1922
QY 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 1982
QY 2186 CGCCGACTTGGTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTATCAA 2245
Db 1983 CGCCGACTTGGTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTATCAA 2042
QY 2246 AGCGGTATGGTGTCTCTCTGAATGGCGGATTTAAACGGCTGTGTCGCCACTCCCA 2305
Db 2043 AGTGGCATGGTGTCTCTCTGAATGGCGGATTTAAACGGCTGTGTCGCCACTCCCA 2102
QY 2306 ACCGGTTTATACCGGAAATGTTTGGGCATACCGGCAAGGGGAAATTTGACACCAAGCAT 2365
Db 2103 GCGGTTTATACCGGAAATGTTTGGGCATACCGGCAAGGGGAAATTTGACACCAAGCAT 2162
QY 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTGTCTCAAGCG 2182

RESULT 6
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
XX
AC
AC
XX

DT 07-OCT-2004 (first entry)
XX H. bizzozeronii urease gene cluster, ureABIEFGH.
XX Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureABIEFGH; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX Helicobacter bizzozeronii.
XX OS US2004142343-A1.
XX PN 22-JUL-2004.
XX PD 12-AUG-2003; 2003US-00639273.
XX PF 16-AUG-2002; 2002US-0404337P.
XX PR (CHAN/) CHANG Y.
XX PA (SIMP/) SIMPSON K W.
XX PA (ZHU/) ZHU J.
XX PI Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
XX DR GENBANK; AF330621.
XX PS Novel isolated nucleic acid molecule having urease gene cluster, and
XX Claim 2; SEQ ID NO 1; 40pp; English.
CC The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureE, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents the H. bizzozeronii
CC urease gene cluster, ureABIEFGH.
XX Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
QY Query Match 41.1%; Score 1184.4; DB 12; Length 8407;
Db Best Local Similarity 68.5%; Pred. No. 1.5e-255; Indels 32; Gaps 2;
Matches 1681; Conservative 0; Mismatches 741;
QY 183 AAATTAAACACAGAGTAGTAATAGTGAAACTCACACCACCAAGAGCAAGAAAGTCTTGT 242
Db 2319 AAATTGGTAGAGGAGTTTAGATGAATTAACCCCTAAAGAGCTGACAGCTCATGT 2378
QY 243 TATATTATCGGGCGAAGTGGCTAGAAAGCGCAAGAGGGCTTAAAGTCAACCAAC 302
Db 2379 TGCATTATCGGGCGAATTTGGCTTAAAAACCGCAAGCAAAATGGGTTAAGCTAATATA 2438
QY 303 CGAAGCCATTGCTTACATTAGTCCCATATTATGACGAGCGCGCGGAAAAAAA 362
Db 2439 CTGAGGCAGTAGCCCTCATCAGTGCCTCATGTGATGGAAGAGCCCGGTCAGGTAATAA 2498
QY 363 CGGTTGCCCGAGCTTATGGAAGAGTGCATGCACCTTTTGTAAAAAAGATGAAGTATGCCCG 422

Db 2499 GTGTGGCGGATTTGATGCAAGAGCGGACGACACTTCTTAAAGCTGATGATGTCATGCCCG 2558
Qy 423 GGGTGGTAAATATGGTTCCTCGATGTAGGTGTAGAGCCACCTTTCCTGTATGGTACGAAC 482
Db 2559 GTGTAGCCCATATGATTCACGAAGTGGGATTTGAAGCTAACTTCCCTGTATGGGACAAAC 2618
Qy 483 TTGTAACTGTGAATTTGGCCCATCGAACACGATGAGCACTTCAAAGCGGGGAAGTGAAT 542
Db 2619 TGGTAACCATTCATACCCCGTGTGAAGATGGTGGGCATAAATTTGGCTCCGGGTGAAGTGA 2678
Qy 543 TTGGTTGGCATAAAGACATCGAGCTCAATGTCAGGCAAGAAAGTAAACCGAACTTGAAGTTA 602
Db 2679 TTTTGAAGAAACGAGACATCACTTTGAAATGACGAGCAACAGGCCACCACTTTAGAAGTGC 2738
Qy 603 CTAATGAAGGGCCATAATCTTGCATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTAAAC 662
Db 2739 ATAAACAAAGCGATCGCCCGTGAAGTGGCTCCCACTTCCACTTCTTTTGAAGTGAATA 2798
Qy 663 AGGCATAAAATTCGATCGTGAAGAAAGCCATATGGCAACGCCCTAGATATTCCTCTGGCA 722
Db 2799 AGCTTTTGAATTTGATCGTGAAGAAAGCCCTATGGCAACGCCCTAGACATTTGCTTCTGGAA 2858
Qy 723 ACAGCTACGCTATGGGCGAGCAACACCGCAAGTGCAGTTGATTTCTCTTGGTGGCA 782
Db 2859 CCGCTGTGGCTTTTGAACCCGGTGAGAAAGAAACCGTGGHAATTTGATTCAAATTTGGCGGTA 2918
Qy 783 GTAAAAAGTGATTTGCATGAACCGGCTTGTGAATAACATCGCGGATGAACGCGCATAAAC 842
Db 2919 ACCAAGCATTTACGCTTTAACTCTCTTGTGGATCGCCAAAGCCGATACTGATGGCAAA 2978
Qy 843 ATAAAGCGTTGCAAGGGGAAATCTCACGGATTT----- 877
Db 2979 AACTTGTCTCTCAACCGCGCAAGAACATGGCTTTTGGTGTGTGAATTTGGCGTGTGGATA 3038
Qy 878 ----ATCAAGTAGAGAGCTCCCATGAATAATGAAGAA--ACAAGAAATATGTAATACCT 930
Db 3039 AAAAAATGAAGAAAGCAATCCGATGAAAAAATCTCTGAAAGAAATATGTTTCTATGT 3098
Qy 931 ACGGACCCCAAGGCGATAAAGTCGCTTAGAGATACCGATCTTTTGGCGAGAGTAG 990
Db 3099 ATGGAACCACTAGCGGCGATAAGTGAATTTGGCGATACCGACCTGATCTTAGAGTCG 3158
Qy 991 AACATGACTATACCACTATGCGGAGAACTTTAAATTTGGCGGGGTAAACTATCCGTG 1050
Db 3159 AACATGACTACCACTTATGCGGAGAAATTAAGTTTGGTGGCGGTAAACCACTTCGCG 3218
Qy 1051 AGGGTATGGTACAGCAATAGCCCTGATGAACACCCCTAGATTTAGTCACTAACG 1110
Db 3219 ATGGGATGGCAAAACCAACAGCCCCAGCAGCCAGCAACTCGATCTTGTGCTCACTAACG 3278
Qy 1111 CGATGATTTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAGCGCAAAA 1170
Db 3279 CCCTGATCGTGGATTTACACCGGCAATTTTAAAGCCGATATTTGGCATTTAAAGTTGGCAAAA 3338
Qy 1171 TCATGSCATTTGGCAAGCAGGCAATAAGACATGCAAGATGGCGTTTGGCAACAATCTTT 3398
Db 1231 TCGTGGGTGGGACAGAGACATAGCAGGGGAGGTATGATTTATACCGCTGGGGAA 1290
Qy 3399 CGGTGGGCCCTGCTACTGAGGCTTTTGGCCGTGAAGGGCTGATTTTACAGCTGGTGGGA 3458
Db 1291 TCGATTTACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCCAATGGCG 1350
Qy 3459 TTGACACCCACATCCATTTATTTCTCCCAACAAATCCCAACAGATTTTGGCAGCGGGA 3518
Db 1351 TTACAAACCATTTTGAAGGGGACAGGTCTGTGTAGATGGCAGCAAGATTCGCACTACTATCA 1410
Qy 3519 TCACAACCATGATTTGGTGGGGACAGGTCCAGCTGATGGACTTAACCGCACTACCATCA 3578
Db 1411 CTCGGGCAAAATGGAACCTTGACCGCATGTTGGCGCAGCAGAAAGATTTCTATGATG 1470
Qy 3579 CTCGGGGCGGTGGAACTTTAAACCATCTCTCCGTGGCTCTTGAAGAAATATGCCATGA 3638

Qy 1471 TGGGCTTTTGGGCAAGGCAATAGCTCTTAGACAAAAAACAACTTGTAGAACAAAGTAGAAG 1530
Db 3639 TGGGCTATTTGGTAAAGGGAATGTGTCTTTATGAACCTCCCTGGTGCATCAACTCGAAG 3698
Qy 1531 CGGGCCGATTTGGTTTAAATTCATGAAGACTGGGSCAACAAACCAAGTGCAGTCCGATC 1590
Db 3699 CTGAGCCATTTGGCTTTAAATCCAGAACTGGGTAGCACACCTTGAGCCATCTACC 3758
Qy 1591 ACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCA 1650
Db 3759 ATTGCTTGAATTTGGCTTGACAAATACGATGTGCAAGTGGCTATCCAACCGATACCTTGA 3818
Qy 1651 ATGAGGAGGTATGTAGATGACACCTTAAATGCAATGAACGGCGGCCCATTCATGCT 1710
Db 3819 ATGAAGCGGCTGTGTGGAAGACACTTTTGAAGCCATTTGCTGGGCGCACTATCCACACTT 3878
Qy 1711 ACCAATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGAGC 1770
Db 3879 TCCACATGAAAGTGTGTGGGCGGACCGCTCCGATGTCATTAAGATGCTTGGCGAAT 3938
Qy 1771 TCAATATTTACCTCTCTCCACACCCCACTATTTCCCTATPACCATTAATACGTTGCGAG 1830
Db 3939 TTAACATCTCTCCAGCTTTACCAACCCCACTTCTTTTCCCGTGAATACAGAACCG 3998
Qy 1831 AACATTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGCGAGGATTTAC 1890
Db 3999 AACACATGGACATGTTGATGTTGCCACCACTTTGGATATAAAACATCAAAGAAATGTCC 4058
Qy 1891 AATTTCTCAAGCCGATCCGCCCGCTCTATCGCGGCTGAAGATGTGCTCCATGATA 1950
Db 4059 AGTTTGTGATTTAGAGATTCGGCCCCCAACCATTCGCCGCTGAGGACAACTCCACGATA 4118
Qy 1951 TGGGTGTGATCGCGATGACAAAGCTTCGGAATTCGCAAGCAATGGGGCGTGCAGCGAAGTGA 2010
Db 4119 TGGGATTTTCTCTATCACCACTCTGACTCCCAAGCGATGCGCGTGTAGGCGAGGTCA 4178
Qy 2011 TTCTCTGAACTTGGCAGACTGCGGATAAGATAAAAAAGAAATTTTGGTAAGCTTCTCTAAG 2070
Db 4179 TCACCCGCACTTGGCAACAGCGACAAAAACAAAAAGAAATTTGGTTCGCTTGCCTGAGG 4238
Qy 2071 ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTACTCAACC 2130
Db 4239 AAAAAAGCGATAATGACAACTTCGCAATCAAGCGCTACATTTCCAAATACACCATCAACC 4298
Qy 2131 CCGCTTTGACCCACCGGCTGAGGAGTATACGCTCTGTGGAAGAGGCGCAAGTCGCG 2190
Db 4299 CCGCTATGACACCGGCAATTTCTGAATATGTGCGCTCTGTAGAAAGTGGGCAATTTGCGCG 4358
Qy 2191 ACTTGTGGTGTGGAATCTTCTGCTTTTGGCGTAAACCCCAAAATCGTGAATCAAAGCG 2250
Db 4359 ATTTGGTGTCTTGGAGTCTTCTTGGCAATTAACCCCAACATGATCATCAAAGCG 4418
Qy 2251 GTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACCGCTCTGTGCCCATCCCGAACCGG 2310
Db 4419 GATTCATCGCACTTCTCTCAATGGGCGATGCAATTCGCTCTATCCCACTCCCGAACCGG 4478
Qy 2311 TTTATTAACCGGAAATGTTTGGGCGATCACGCAAGGCGAAATTTTGCACACCGACATCACTT 2370
Db 4479 TGTATTACCGGAAATGTTTGGCCCAACATGTTGAAGCCAAATTTGACACCAATATCACTT 4538
Qy 2371 TTGTTTCCAAGTCCCTCTATGAAATGGCGGTGAAGAAAGGCTGGGCTTTAGAGCCCAAG 2430
Db 4539 TTGATCCCAAGTGGCTTATGACAAAGCGCATTTAAAGAAAGTGGTGGCTTGAAGAGTGG 4598
Qy 2431 TTCTACCGGTCAAAAACTGGCGTAACTCAACAGAAAGACTTCAAGTTCAACGACAAAA 2490
Db 4599 TTTTGGCAAGTTAAAACTGGCGCAACATCAACAAAAAGACCTCAAAATTCACGATGTTA 4658
Qy 2491 CGGCAAAATACCGTCGATCCGAAACCTTTCGAGGTCTTTGTAGATGCGCAAACTCTGCA 2550
Db 4659 CCGCACATCGAAGTCAATCTGAAACCTTAAAGTTAAAGTGGATGCAAGAGGTTA 4718

QY 2551 CCTCTAAACCCACTCGCAGTGCCTCTAGCCAGCGCTACACTTCTCTTAGG 2604
|||||
Db 4719 CTTCCAAAGCAGCGGATAAAATCAGCTAGCACAACTCTACAACTTGTCTTAGG 4772
|||||

RESULT 7

ID AAT44351 standard; cDNA; 2735 BP.
XX AAT44351;

XX AC AAT44351;

DT 16-OCT-2003 (revised)

DT 11-FEB-1997 (first entry)

DE H. pylori ureA + ureB gene locus in pORV214.

XX Urease; ureA gene; ureB gene; vaccine; ds.

XX Helicobacter; strain CPM630.

OS unidentified bacteriophage; T7.

OS Chimeric.

PH Key Location/Qualifiers

FT promoter 1..16

FT /tag= a

FT /note= "T7 promoter provides transcription initiation for

FT the urease genes"

FT 33..43

FT /tag= b

FT /function= "operator"

FT /note= "Lac operator provides inducible expression of the

FT complement (46..67)

FT /tag= c

FT /note= "BL1 primer"

FT 102..818

FT /tag= d

FT /product= "urease A subunit"

FT 822..2531

FT /tag= e

FT /product= "urease B subunit"

FT 2546..2569

FT /tag= f

FT /note= "BL2 primer"

FT 2693..2735

FT /tag= g

FT /note= "T7 terminator"

XX WO9633732-A1.

PN 31-OCT-1996.

PD 25-APR-1996; 96WO-US005800.

XX 28-APR-1995; 95US-00431041.

PR 06-DEC-1995; 95US-00568122.

XX (ORAV-) ORAVAX INC.

PA Lee CK, Monath TP, Ackerman SK, Thomas WD, Soman G, Kleanthous H;

XX Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;

PI WPI; 1996-497373/49.

DR P-PSDB; AAW07193, AAW07194.

DR Vaccine for inducing mucosal response to Helicobacter - contg. multimeric

XX urease complex and pref. an antibiotic, anti-secretory agent or bismuth

PT salt.

PT Disclosure; Page 68-70; 98pp; English.

XX A cDNA clone (AAT44351), derived from pORV214, includes the ureA and ureB

XX sequences coding for the urease A (AAW07193) and urease B (AAW07194)

CC

CC subunits of Helicobacter pylori clinical isolate CPM630. To obtain
CC pORV214, a genomic DNA library of CPM630 was screened with anti-
CC Helicobacter urease antibody. A 17 kb SalI fragment from an isolated
CC clone was subcloned into pUC18 to give pSCP1. PCR primers (AAT44352-53)
CC were used to amplify a 2.5 kb fragment from pSCP1, which was inserted
CC into pET24+ to give pORV214. The vector was utilised in the prodn. of
CC recombinant, enzymatically inactive, multimeric urease in E. coli
CC transformants for use in vaccines to treat or prevent Helicobacter
CC infection. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 2735 BP; 875 A; 562 C; 635 G; 663 T; 0 U; 0 Other;

Query Match 37.3%; Score 1075.2; DB 2; Length 2735;
Best Local Similarity 66.0%; Pred. No. 3.5e-231;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

QY 182 AAAATTTAACACAGGAGTAATAGGTGAAGTCACTCACACCCCAAGAGCAGAAAGTCTTCTTG 241
|||||
Db 78 ACACCTTTAAGATAGGAGATAGATGAATCACTCACCAAAAGAGTTAGTAAGTTGATG 137
|||||
QY 242 TTATATTATGCGGCGAAGTGGCTAGAAAGCGAAAGCAGAGGCTTTAAAGCTCAACCAA 301
|||||
Db 138 CTCACCTACGCTGAGNAATTGGCTTAAACACGCAAGAAAGGCATTAGCTTAACAT 197
|||||
QY 302 CCCAAGCCATTGCTTACATTAGTGCCCATATTATGACGAAAGCGCCGCTGGAAAAAAA 361
|||||
Db 198 GTAGAAGCAGTAGCTTTGATTAGTGCCCATATTATGGAAGAGCAGAGCTGGTAAAAAG 257
|||||
QY 362 ACCGTGGCCAGCTTATGGAAGAGTGCATGCATCTTTTGAAGAAAGATGAAGTAATGCC 421
|||||
Db 258 ACTGCGCGTGAATTGATGCAAGAGGCGCAGCTCTTTTAAACACCATGATGTGATGGAT 317
|||||
QY 422 GGGTGGGTAAATAGTTCCCGATCTAGTGTAGAGCCACCTTTCTCTGATGGTACGAAA 481
|||||
Db 318 GGGTGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTGATGGACTAAA 377
|||||
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAACCAAGATGAGCACTTCAAAGCGGCGAAGTGA 541
|||||
Db 378 CTGTTAACCGTGCTATACCCCTATTGAGGCCAATGGTAAATAGTTCTCTGGTGGTGG 434
|||||
QY 542 TTTGGTTGGATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAAACCGAACTTGA 601
|||||
Db 435 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTAGCGTGAAGTT 494
|||||
QY 602 ACTAATGAAGGCGCTAAATCTTGCATGTGGGTAGCCATTTCCCATCTCTTTGAGCTAAC 661
|||||
Db 495 AAAAATGTTGGCGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGA 554
|||||
QY 662 AAGGCACATAAAATTCGATCGTGAAGAGCCCTATGGCAAAACGCTAGATATTCCCTCTGC 721
|||||
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTGCTTAAACGCTTAGACATTGCGAGCGG 614
|||||
QY 722 AACACGCTACGCATTGGGCGAGACAAACCCGCAAAAGTGCAGTTGATTCCTTTGGTGGC 781
|||||
Db 615 ACAGCGGTAAAGATTTCAGCGCTGCGGAAGAAAAATCCGTAGAAATTGATTGACATTGGCGGT 674
|||||
QY 782 AGTAAAAAAGTGAATTCGCATGAACGGGCTTGTGTAATAACATCGCGGATGAACGCCATAA 841
|||||
Db 675 AACAGAAGAAATCTTTGGATTTAAACGCAATTTGTTGATAGACAAGCAGACAAACGAAGCAA 734
|||||
QY 842 CATAAAGCGCTTGACAAGGCGAAATCTCAGCGAATTT----- 877
|||||
Db 735 AAAATTTGCTTTACACAGAGCTAAAGAGCGTGGTTTTCATGGCGCTAAAGCGATGACAC 794
|||||
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA---ATGAAAAAACAGAAATATGTAAT 926
|||||
Db 795 TATGTAAAAACAATTAAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGAAATATGTTCT 854
|||||
QY 927 ACCTACGACCCCAAGCGCATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
|||||
Db 855 ATGTATGGTCTACTACAGCGATAAAGTGAGATTGGCGGATACAGACTTTGATCGCTGAA 914
|||||
QY 987 GTAGAACATGACTATATACCCTATGCGAAGAACTTAAATTTTGGCGCGGTAAAACTATC 1046
|||||

```
Db 915 GTAGAACATGACATACACCATTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTA 974
Qy 1047 CGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATCACT 1106
Db 975 AGAAGAGCATGAGCCAACTTACAAACCTTAGCAAGAGTGGATTTAATATCACT 1034
Qy 1107 AACGGATGATATTCGACTACACCGGATTTACAAAGCGACATTTGGGATTTAAACAGCG 1166
Db 1035 AAGCCTTTAATCGTGGATTTACACCGGTATTTATAAAGCGGATATTTGGTATTTAAAGATGC 1094
Qy 1167 ABAATCCATGGCATTCGCAAGGAGGAGAAACAAGGACATGCAAGTGGCGTAAAGCCCTCAT 1226
Db 1095 ABAATCGCTGGCATTTGGTAAAGCGGTAAACAAGACATGCAAGATGGCGTTAAACAAAT 1154
Qy 1227 ATGCTCGTGGGTGGGCAAGACACTAGCAGGGGAAAGGTATGATTTATACCGCTGGG 1286
Db 1155 CTTAGCTAGGTCCTGCTACTGAAGCCTTAGCCGGTGAAGGTTGATCGTAAACGGCTGGT 1214
Qy 1287 GGAATCGATTTACACACCCACTTCTTTCTCCAAACAAATTCCTACCGCTCTAGCCAAT 1346
Db 1215 GGTATTTGACACACATCCACTTCAATTCACCCCAACAAATCCCTACAGCTTTTGCACGC 1274
Qy 1347 GCGTTACACCATGTTTGGAGGGGSCACAGGTCCTGTAGATGSCACGAATGCGACTACT 1406
Db 1275 GGTGTAAACACCATGATTTGGTGGTGGAAACCGGTCCTGCTGATGGCACATAATGCGACTACT 1334
Qy 1407 ATCACTCCGGGCAAAATGGAATTTGCACGCGATGTTGGCGCAGCAGAGAGATTTCTATG 1466
Db 1335 ATCACTCCAGCAGAGAAATTTAAATGATGCTCAGAGGGCTGAAGATATTTCTATG 1394
Qy 1467 AATGTGGGCTTTTGGCAAGGCAATPAGCTCTAGCAAAAACAACTTTAGAACAAAGTA 1526
Db 1395 AATTTAGGTTTCTTGGCTAAAGGTTAACGCTTCTAACGATGCGAGCTTAGCCGATCAAAAT 1454
Qy 1527 GAAGCGGGCGGATGGTTTAAATTCGATGAAGACTGGGGCAACAACCAAGTGGCGATC 1586
Db 1455 GAAGCGGGTGGATTTGGCTTTAAATTCAGAAAGACTGGGGCACCACTCTCTTCTGCAATC 1514
Qy 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATACA 1646
Db 1515 AATCATCGGTTAGATGTTGGGCAAAATACGATGTGCAAGTGCATATCCACAGACACT 1574
Qy 1647 GTCAATGAGGCGGTTATGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAGAAGACACTATGCTGCTATTTGTGGACGCACTATGCAC 1634
Qy 1707 GCCTACACATTTAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGCTGGCGGGGACACGCTCTGATATATTATAAGTAGCCGGT 1694
Qy 1767 GAGCTCAATATTCACCTCTCCACACCCCACTATTCCTATACCATTAATACGGTT 1826
Db 1695 GAACACAACTCTTCCCGCTTCCACTAACCCCACTCCCTTTCCACCGTGAATACAGAA 1754
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGAT 1886
Db 1755 GCAGACACATGGACATGCTTATGTTGTGCCACCACTTGGATTAAGAATTAAGAAGAT 1814
Qy 1887 TTAACAATTTCTCAAAGCGTATCCGCGCCGGCTCTATCGCGGCTGAAGATGTGCTCAT 1946
Db 1815 GTTCAGTTCGCTGATTTCAAGGATCCGCGCTCAAAACCATTTGGCGCTGAAGACACTTTGCTAT 1874
Qy 1947 GATATGGGTGATGCGGATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGGCAA 2006
Db 1875 GACATGGGATTTTCTCAATCACCAGTTCTCACTCTCAAGGCTGAGGGCGGTGGGTGAA 1934
Qy 2007 GTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATTAAGAAATTTGGTAAAGCTTCCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAACAGAAAGAAATTTGGCGGCTTGAA 1994
Qy 2067 GAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATC 2126
|||||
```

```
Db 1995 GAAGAAAAAGCGGATTAACGAACTTCAGGATCAACCGCTACTTGTCTAAATACACCAATT 2054
Qy 2127 AACCCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTAGGTTTCAGTAGAAGTGGGCAAGTG 2114
Qy 2187 GCGGACTTGGTGGTGGGAATCTGCGCTTTTGGCGGTAAACCCAAATCGTGTATCAAA 2246
Db 2115 GCTGACTTGGTATTTGTGGAGTCCAGCAATCTTTGGCGGTGAACCCCAACATGATCATCAA 2174
Qy 2247 GCGGATATGTGGTCTTCTCTGAAATGGGCGATTTCTAACGCGTCTGTGCCCACTCCCCAA 2306
Db 2175 GCGGATTCATTTGGTTAAGCCAAATGGGCGATGCGAACGTTCTATCCCTACCCCAAA 2234
Qy 2307 CCGGTTTATTAACCGCAAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACACCAAGCATC 2366
Db 2235 CCGGTTTATTAACGAGAAATGTTTCGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2294
Qy 2367 ACTTTGTTTCCAAAGTGGCTATGAAATGGGCTGAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2295 ACTTTTGTGTCTAAGCGGCTTATGACAAAGGCAATTAAGAAAGAAATTAGGACTTTGAAAGA 2354
Qy 2427 CAAGTTCCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCACCGAC 2486
Db 2355 CAAGTGTTCGCGTAAAAAAATTCGAGAAATATCACTAAAGAAACATGCAATTCACCGAC 2414
Qy 2487 AAAACGGCAAAAAATCACCGTCGATCCGAAACCTTCGAGGTCCTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTTGAAGTCAATCCTGMAACTTTACCATGTGTTGCTGATGGCAAGAA 2474
Qy 2547 TGCACCTCTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 2475 GTAACCTCTAAACAGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCAATTTCTTAGG 2532

RESULT 8
AAQ90180
ID AAQ90180 standard; DNA; 2619 BP.
XX
AC AAQ90180;
XX
AC AAQ90180;
XX
DT 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX
XX Helicobacter felis urease ureA/ureB operon.
XX
KW Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein; ss.
XX
OS Helicobacter felis.
XX
FH Location/Qualifiers
RBS 31..37
FT /*tag= b
FT /note= "ureA Shine-Dalgarno site"
FT CDS 43..756
FT /*tag= a
FT /EC_number= "3.5.1.5"
FT FT 756..759
FT /note= "UreA"
FT RBS 766..2475
FT /note= "ureB Shine-Dalgarno site"
FT CDS 766..2475
FT /*tag= c
FT /EC_number= "3.5.1.5"
FT /note= "UreB"
XX
WO9514093-A1.
XX
XX 26-MAY-1995.
XX
XX 19-NOV-1993; 93WO-EP003259.
XX
```



```
Db 1706 ACATGACATGTTAATGGTGGCCACCTGTGATAAAAGTATCAAGGAAGATGTGAGT 1765
Qy 1894 TTTCTCAAAAGCCGTATCCGCGCGGTCTATCCGCGCTGAAGATGTCTCCATGATGG 1953
Db 1766 TTGCCGATTCGAGGATTCGCCCCCAACTATCCGCGCTGAGACCAACTCCATGATGG 1825
Qy 1954 GTGTGATCGCGATGAAGAAGTCGGATTCGCAAGCAATGGGGCTGAGGGGCAAGTATTC 2013
Db 1826 GGATCTTTTCTATCAGCAGTCGCACTCTCAGGCTATGGGACGCTAGGCGAGTGTATCA 1885
Qy 2014 CTCGAATCTGGCAGATCGCGATGAAGATGAAGAAAGAAATTTGTAAGCTTCCTGAAGTG 2073
Db 1886 CAGCGACTTGGCAGACAGCAGACAAACAAACAAAGAGATTGGCGCTTGAAGAGGAA 1945
Qy 2074 GCAAGATTAACGATAATTTCCGATTAAGCGCTACATCTCAAAATACACTATCAACCCCG 2133
Db 1946 AAGCGGATAACGACACTTCCGATCAAGCGCTACATCTCTAATATACACCATCAACCCG 2005
Qy 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACT 2193
Db 2006 GGATCGCGATGGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAATACGCCGACC 2065
Qy 2194 TGGTGTGTGGAATCTGCGCTTTTGGCGTAAACCCAAATCGTGATCAAGGGCGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGGCTTTCTTTGGCATTAAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATGGGGGATTTAAACGCTGTGCGCACCTGCCCAACCGGTTT 2313
Db 2126 TTAATGGCTCTCTCAATGGGGGATGCCAATGGCTATTTCCACCCCTCAGCGCGTCT 2185
Qy 2314 ATTACCGCGAAATGTTTGGGCATCACGGCAAGCGGAAATTTGACACGAGATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATCGGGAACAAATTCGACACCAATATCACTTTG 2245
Qy 2374 TTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAGCTGGGCTAGAGCGGCAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTCAAGCGAGGGATCAAGAAGAACTAGGGCTAGATCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAATGCGGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAACGG 2493
Db 2306 CGCAGTGAATAACTGTGCAATATCACTAAAGAGACCTCAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTT 2553
Db 2366 CACATATGATGTCAACCTGAAACCTTATAAGTGAAGTGAAGTGGATGGCAAGAGTAACTT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTTAGCCAGCGCTACACTTTTCTTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATGAGCGCTAGCGCAACTTTATAATTTGTTCTAGG 2476
```

RESULT 9

```
AAAT45680
ID AAAT45680 standard; DNA; 2619 BP.
XX
AC AAAT45680;
XX
DT 16-OCT-2003 (revised)
DT 07-MAR-1997 (first entry)
XX
XX H. felis structural urease gene.
XX
KW Urease; UreA; UreB; heat shock protein A; HspA; vaccine; immunogen;
KW antigen; antibody; chronic gastritis; ulcer; pIL205; ss.
XX
OS Helicobacter felis; strain ATCC 49179.
XX
XX Key Location/Qualifiers
XX RBS 31..36
XX FT /*tag= a
XX FT 43..753
XX FT /*tag= b
XX
```

/standard_name= "URE A"
/product= "urease subunit A"
756..759
/*tag= c
766..2475
/*tag= d
/standard_name= "URE B"
/product= "urease B subunit"
WO9634624-A1.

07-NOV-1996. 96WO-EP001834.
02-MAY-1996; 95US-00432697.
19-MAY-1995; 95US-00447177.
XX (INSP) INST PASTEUR.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX Labigne A, Suerbaum S, Ferrero RL, Thiberge J;
XX WPI; 1996-505900/50.
XX P-PSDB; AAW06729, AAW06730.
XX New immunogenic compsn. contg. UreB and HspA antigens of Helicobacter -
XX for treatment and prevention of esp. H pylori infection, also new
XX antibodies specific for these antigens.
XX Example 1; Page 96-100; 184pp; English.

CC A portion (AAAT45680) of plasmid pIL205 (NCIM I-1355) comprises the
CC Helicobacter felis structural urease gene encoding urease subunits A
CC (AAW06729) and B (AAW06730). pIL205 was obtd. by examining the
CC expression of urease activity by H. felis cosmid clones and
CC identification of the H. felis genes required for urease expression when
CC cloned in E. coli. The urease gene can be used to design probes for the
CC detection of Helicobacter infection, or utilized in the prodn. of
CC recombinant urease subunits for use in novel immunogenic compsns.
CC (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;
Query Match 36.9%; Score 1062.6; DB 2; Length 2619;
Best Local Similarity 65.5%; Pred. No. 2.3e-238;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;

Qy 192 ACAAGGAGTAAATAGTGAAGCTCACCCAAAGAGCAAGAAAGTTCTTTTATATATG 251
Db 29 ATAAGGAGTTAGGATGAAGCTAAACGCTAAAGAACTAGACAAGTAAATGCTCCATTATG 88

Qy 252 CGGCGAAGTGGCTAGAAAGCGCAAGCGGCTTAAAGCTCAACCAACCGGAAGCCA 311
Db 89 CGGCGAATTGGCAGAAAGACGCTTGGCGCTGTGTGAAACTCAATTACACCGAAGCGG 148

Qy 312 TTGCTTACATTAGTCCCATATTATGACGAAGCGCGCTGGAAAAAACCCTTTGCC 371
Db 149 TCGCGCTCATTTAGCGGGCTGTGATGAAGAGCGCGTGTGATTAATAAGCGTGGCGG 208

Qy 372 AGCTTATGGAAGAGTGCATGCATCTTTTGAAGAAAGATGAAGTAAATGCCCGGGTGGTA 431
Db 209 ATTTGATGCAAGAGGCGAGGACTTTGGCTTAAAAAAGAAATGTGATGACGCGTAGCAA 268

Qy 432 ATATGTTCCCGATCTAGGTGTAGAAGCCACTTTCTGTAGTGGTACGAACTTTGAACGTG 491
Db 269 GCATGATTCATGAAGTGGGATTTGAAGCTAACTTCCCGATGAAGCAACCAAGCTTTGAAC 328

Qy 492 TGAATTGGCCCATCGAACACGATGAGCACTTCAAGAGCGGCGAAGTGAATTTGGTTGG 551
Db 329 TCCACACTCCGATAGAGGATTAATGGCAATTAGCCCGCGGAGGTCTTCTTAA---NA 385

Qy 552 ATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAACCGAATTGAGGTTACTAATGAAG 611

Db 386 ATGAGGACATTACTATTAACGCGCGCAAGAGCCATTAGCTTTGAAGTGAATAATAAAG 445
Qy 612 GGCCTAAATCCTTGTGATGTTGGGTAGGCATTTCCACATTTCTTTGAAGCTAAACAGGCACTAA 671
Db 446 GCGATCGTCTGTGAGGTGGGATACATTTCCACTTTCTTCGAAATGAATAGCTCTTGG 505
Qy 672 AATTGCGATCGTGAAGAGCTTAGGCAACCGCTAGATATTCCTCTGGCAACAGCCTAC 731
Db 506 ACTTCGATCGCGCAAAAGCTTTTGAACAGCGCTAGACATTCGGAACAGCGGTGC 565
Qy 732 GCATTTGGGCGAGCAAAACCGCAAGTGCAGTTGATTCCTCTTGTGTGCGAGTAAAGAG 791
Db 566 GCTTTGAACCCGGGAGGAGAAAGTGTGGAACCTATTGACATCGCGCGGGAATAAGCGCA 625
Qy 792 TGAATGGCATGAACGGCTTGTGAATAACATCGCGGATGAACGCCATAAACAATAAGCGC 851
Db 626 TCTATGGCTTTAATCTTTTGGTGGATCGCAAGCCGATCCGATGGTAAAAAATCTCGGCT 685
Qy 852 TTGACAAAGCGCAATCTCACGGATTT----- 877
Db 686 TAAACCGCGCTAAGAGAAAGGTTTGGGTCTGTAAACTCGGTTGTGAAGCGACTAAAG 745
Qy 878 -----ATCAAGTAAGGAGACTCCCATGAAATGAAAAAACAAGATATGTAATACCTACG 933
Db 746 ATAAACNATAAGGAAAAAACCATGAAAAAGATTTCACGAAAAAGATATGTTTCTATGTATG 805
Qy 934 GACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC 993
Db 806 GTCCCACTACCGGGGATCGTGTAGACTCGGCGACACTGATTTGATCTTTAGAGATGGAGC 865
Qy 994 ATGACTATACCACTATATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGG 1053
Db 866 ATGATTGACCACTTATGTGTGAAGATCAAAATTTGGGGCGGTAAAACTATCCGTGATG 925
Qy 1054 GTATGGGTGAGAGCAATAGCCCTGATGAAAAACAACCTAGATTAGTCACTAAACGCGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTGTGGTCTCACTAACGCC 985
Qy 1114 TGAATTATGACTACACCGGGATTTCAAGCGGACATTTGGGATTTAAAAACGCAAAATCC 1173
Db 986 TCAITGTGACTATAGGGCAATTTCAAGCGGACATTTGGGATTTAAAGCGGCAAGATTG 1045
Qy 1174 ATGGCATTTGGCAAGCAGCAACAGGACATGCAAGTGGCTTAAGCCCTCATATGGTGC 1233
Db 1046 CAGGCATTTGGCAAGGAGGCAATAAGGACATGCAAGTGGCTGAGTAATAATAATCTTTGGC 1105
Qy 1234 TGGGTGTGGGCACAGAACACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCTGCTACAGAGCTTTGGCAGCTGAGGGCTTGATTGTAAACCGCTGGTGGCATCG 1165
Qy 1294 ATTCAACACCCACTTCTTTCTCCAACAATTCCTTACCGCTTAGCCAATGGCGTTA 1353
Db 1166 ATACGCATATTCACTTTATCTCTCCCAACAAATCCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCATGTTTGGAGCGGCAAGGCTCTGTAGATGGCAAGATGGCACTATATCACTC 1413
Db 1226 CAACCATGATTGGAGGAGGCAAGGACCTTGGGATGGCAAGATGGCACTATCACTC 1285
Qy 1414 CGGCAATATGGAATTTGCAACCGCATGTTTGGCGCGCAGCAAGAGTATTCTATGAATGTGC 1473
Db 1286 CCGGAGCGGCTAATCTAAAGATATGTTGCTGTGCGCGGAGGAGATGCGCATGAATCTAG 1345
Qy 1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGAGAACAAAGTAGAGCGG 1533
Db 1346 GCTTTTGGCTTAAGGGGAATGTCTTACGAACCTCTTTACCGCATCAGATTGAAGCAG 1405
Qy 1534 GCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGGGATCGATCACT 1593
Db 1406 GGGCGATTGGTTTAAATTTCCAGAGACTGGGAGGCAACCTCGCAGCTATTTCCCACT 1465
Qy 1594 GCTTCAGCGTGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATG 1653

Db 1466 GCCTCAATGTGCGCGATGAATACGATGTGCAAGTGGGTATCCACACCGATACCCCTTAACG 1525
Qy 1654 AGGCAGCTTATGTAGATGACACCCCTAAATGAACGGGCGGCCATTCATCGCTACC 1713
Db 1526 AGCGGGCTGTGTAGAAGACACCCCTAGAGGGATTTGCGGGCGCACCATCCATACCTTCC 1585
Qy 1714 ACATTAAGGAGCGGTGGAGGACACTCACCTGANTTTATCACCATGGCAGCGAGCTCA 1773
Db 1586 ACATTAAGGGGCTGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTTCTACCTCTCCACACCCCGCTATTCCTCTATACCATTAATTAACGTTTCAGAAC 1833
Db 1646 ACATTTCTACCCGCTCTACTAACCCGACCATCTCTTTTACCAAAACACTGAAGCCGAGC 1705
Qy 1834 ACTTAGACATGCTCATGACATGCCACCATAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACACTTGTGATAAAGTATCAAGGAAGATGTGCAGT 1765
Qy 1894 TTTTCTCAAAAGCCGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATGG 1953
Db 1766 TTGCCGATTCGAGGATTCGCCCCCAAACTATTCGGGCTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCCGATGACAAAGCTCCGATTCGCAAGCAATGGGCGTGCAGGCGAAGTATTC 2013
Db 1826 GGATCTTTTCTATCACCGACTCCGACTCTCAGGCTATGGGACGCTAGGCGAGGTATCA 1885
Qy 2014 CTCGAATTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCTCTCAAGATG 2073
Db 1886 CAGGCACTTTGGCAGACAGCAGACAAAAAAGAGTTTGGCGCTTGAAGAGGAA 1945
Qy 2074 GCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGCGATTAACGACAACTTCGCGATCAACCGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGAACCCAGCGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGATCCCGACT 2193
Db 2006 GGATCGCGATGGGATTTCTGACTATCTGCGCTCTGTGGAAGTGGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGTGTGGAATCTCTGCTTTTGGCGTAAACCCAAATCGTGATCAAGAGCGGTA 2253
Db 2066 TCGTGTCTTTGGAGTCCGGCTTTCTTTGGCAATTAAGCCCAATATGATTATTAAAGGGCGAT 2125
Qy 2254 TGGTGTCTTCTCTGAATGGGCGATTTCTAAACGCTCTGTGCCACTCTCCCAACCGGTTT 2313
Db 2126 TTATTTGGCTCTCTCAATGGGCGATGCCAATGGCTCTATTTCCACCCCTCAGGCCGTCT 2185
Qy 2314 ATTACCGCAAAATGTTTGGGCATCAACGGAAGGAAATTTGACACAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAAACAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGGCAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTCAAGGCGAGGATCAAGAAAGAACTAGGGCTAGATTCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAACCTCCGTTAAACATCAACAAAGAAAGACTTCAAGTTTCAACGACAAAAACGG 2493
Db 2306 CGCGAGTGAAGAACTCTGCGAATATCACTAAAAGGACCTCAATTCACAGATGTGACCG 2365
Qy 2494 CAAAAATCAACCGTTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTTCGACCT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTATAAGGTGAAGTGGATGGCAAGAGTAACCT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTCTTAGCCGAGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTTAGCCTTAGCGCAACTTTATATATTTTCTTAGG 2476

RESULT 10
AAQ75319
ID AAQ75319 standard; DNA; 2619 BP.
XX
AC
XX AAQ75319;

DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 22-JUN-1995 (first entry)
 XX Urease ureA/ureB gene.
 XX
 XX
 XX Urease; ureA gene; ureB gene; immunogen; vaccine; diagnostic;
 KW Helicobacter pylori; heat shock protein; HSP; chaperonin; pili205; ss.
 XX
 OS Helicobacter felis; ATCC 49179.
 XX
 XX
 PH Key Location/Qualifiers
 FT RBS 31..36
 FT FT /*tag= a
 FT CDS 43..756
 FT FT /*tag= b
 FT FT /label= ureA gene
 FT FT /note= "encodes urease A subunit"
 FT FT 756..759
 FT FT /*tag= c
 FT CDS 766..2475
 FT FT /*tag= d
 FT FT /label= ureB gene
 FT FT /note= "encodes urease B subunit"
 XX
 PN W09426901-A1.
 XX
 XX
 XX 24-NOV-1994. 94WO-EP001625.
 XX
 XX 19-MAY-1994; 94WO-EP001625.
 XX
 XX 19-MAY-1993; 93EP-00401309.
 XX 19-NOV-1993; 93WO-EP003259.
 XX
 XX (INSP) INST PASTEUR.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 XX
 XX
 PI Labigne A, Suerbaum S, Ferrero R, Thiberge J;
 XX
 XX WPI; 1995-006797/01.
 DR P-PSDB; AAR67371.
 DR
 XX DNA from Helicobacter pylori and Helicobacter felis - used to develop
 PT prods. for detection, treatment and prevention of Helicobacter infection.
 PT
 XX Disclosure; Fig 3; 168pp; English.
 PS
 XX Vaccine compositions include the A and B subunits (given in AAR67371) of
 CC H. felis urease encoded by the ureA/ureB gene (AAQ75319) region of the
 CC urease gene cluster of pili205 (NCNM I-1355), as well as the heat shock
 CC proteins HSPA (AAR67374) and HSPB (AAR67373) encoded by the urease-
 CC associated HSP gene cluster region (AAQ75321) of pili689 (NCNM I-1356).
 CC Recombinant products are expressed in Escherichia coli. (Updated on 25-
 CC MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 2619 BP; 773 A; 570 C; 664 G; 612 T; 0 U; 0 Other;
 Query Match 36.8%; Score 1061; DB 2; Length 2619;
 Best Local Similarity 65.5%; Pred. No. 5.4e-228;
 Matches 1605; Conservative 0; Mismatches 805; Indels 41; Gaps 2;
 QY 192 ACAAGGAGTAATAGGTCAAACTCACCCCAAGAGCAAGAAAGTCTCTGTATATTATG 251
 DB 29 ATAGGAGTTAGGATGAAACTAACCGCTTAAGAACTAGACAGTTAATGCTCCATTATG 88
 QY 252 CGGCGAAGTGGCTAGAAAGCGCAAGACAGAGGGCTTTAAAGCTCAACCAACCCGAAGCCA 311
 DB 89 CGGCGAGATTGGCAGAGAGAGCGCTTGGCGGTGGTGTGTAACCTCAATTACACCGAAGGG 148
 QY 312 TTGCTTACATTAGTGGCCATATTATGACGAGCGCGCGTGGAAAAAACCCTTGCC 371
 DB 149 TCGCGCTCATTTAGCGGGCGTGTGATGGAAGAGCGCGTGTGTTAATAAAGCGTGGCGG 208
 QY 372 AGCTTATGGAAGAGTGCATGCACACTTTTGGAAAAAGATGAAGTAAATCCCGGGGTGGGTA 431
 DB 209 ATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAGAAAAATGTGATGCACGGCTAGCAA 268
 QY 432 ATATGTTCCCGATCTAGGTGTAGAAGCCACTTCTCTGATGGTACGAAACTTTGTAATG 491
 DB 269 GCATGATTCATGAAGTGGGATTAAGCTAACTTCCCGGATGGAAACCAAGCTTGTAACTA 328
 QY 492 TGAATTGGCCCATCGAACACAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGCG 551
 DB 329 TCCACACTCGGTAGAGATAATGGCAAAATTAGCCCGCGGAGGTCTTCTTAA--AA 385
 QY 552 ATAAAGACATCGAGCTCAATGCAAGGCAAGAGTAACCGAATTGAGGTTACTAATGAAG 611
 DB 386 ATGAGGACATTACTATTAAACCGCGCAAGAGCAATTAGCTTGAAGTGAATAAAG 445
 QY 612 GGCTAAATCCTTGCATGTGGTAGCCATTTCACATCTTTTGAAGCTTAACGAAGGCACTAA 671
 DB 446 GCGATCGTCTGTGAGGTGGGATCACATTTCCACTTCTTCGAAGTGAATAAGCTCTTGG 505
 QY 672 AATTTCGATCGTGAAGAGCCTATGCGCAAGCGCTTAGATATTCCCTCTGGCAACACGCTAC 731
 DB 506 ACTTCGATCGGCNAAAAGCTTTTGCNAAACGCTTAGACATTCGATCTGGAACAGCGGTGC 565
 QY 732 GCATTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTAAAAAG 791
 DB 566 GCTTTGAACCCGCGGAGGAAAAAGTGTGGAACATCAATTGACATCGCGGGGAATAAGCGCA 625
 QY 792 TGATTGGCATGAACGGGCTTGTGATAAACATCGCGGATGAACGCCATAAACATAAAGCGC 851
 DB 626 TCTATGGCTTTAAATCTTTGGTGGATCGCAAGCCGATCGGATGGTAAAAAACTCGGCT 685
 QY 852 TTGACAAAGCGAAATCTCACGGATT----- 877
 DB 686 TAAAAACCGCTTAAAGAAAAAGTTTGGGTCTGTAACTCGCGTGTGTGAACGCACTAAG 745
 QY 878 ----ATCAAGTAAGGAGACTCCCATGAAAAATGAAAAAACAAGAAATATGTAATACCTACG 933
 DB 746 ATAAACAATAAGGAAAAACCATGAAAAAGATTTACGAAAAAGAAATATGTTCTATGTATG 805
 QY 934 GACCCACAAAGCGGATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAAGTAGAAC 993
 DB 806 GTCCCACTACCGGGGATCGTGTTAGACTCGCGCACTGATTTGATCTTAGAAGTGGAGC 865
 QY 994 ATGACTATACCACTATGSGGAAGAACTTAAATTTGGCGGGTAAACATCTCGTGAAG 1053
 DB 866 ATGATTGCACCACTTATGGTGAAGATCAAAATTTGGGCGGTAAAACTATCCGTGATG 925
 QY 1054 GTATGGTTCAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTCACTAACCGGA 1113
 DB 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTGGTGTCTCACTAACGCC 985
 QY 1114 TGATTATCGACTACACCGGGATTTAAAGCCGACATTTGGGATTTAAACCGGCAAAATCC 1173
 DB 986 TCATTGTGGACTATACGGGCATTTACAAAGCCGACATTTGGGATTTAAAGACGCGCAAGATTG 1045
 QY 1174 ATGCGATTGGCAGGCAAGGAAACAGGACATGCAAGATGGCGTAAGCCCTCATATGTCG 1233
 DB 1046 CAGGCATTTGGCAAGCGGAGGCAATAAGGACATGCAAGATGGCGTAGATAAATAATCTTTCG 1105
 QY 1234 TGGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAATCG 1293
 DB 1106 TAGTCTCTGTACAGAGCTTTGGCAGCTGAGGCTTGATTGTAAACCGCTGTTGTCATCG 1165
 QY 1294 ATTACACACACCACTTCTTTCTCCACAAATTTCCCTACCGCTCTAGGCCAATGGCGTTA 1353
 DB 1166 ATACGCATATTCACTTTATCTCTCCCAACAAATCTCTACTGCTTTTCCAGCGGGGTTA 1225
 QY 1354 CAACCATTTTGGAGCGGCGACAGTCTGTAGATGGCAGCAATGCGCACTACTATCACTC 1413
 DB 1226 CAACCATGATTTGAGGAGGCGACAGGACTCTCGGATGGCAGCAATGCGCACTACTCACTC 1285

Qy	1414	CGGCAAAATCGAACTTTCGACCGGATGTTGGCGGACAGAGAAGTAGTATTCCTATGAATGTGG	14173
Db	1286		1345
		CCGACCGCGCTAAATCTAAAAAAGTATGTTGGCTGCAGCCGAAGAATA CGCCATGAATCTAG	
Qy	1474	GCTTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAGTAGAGCGG	1533
Db	1346		1405
		GCTTTTTGGCTAAGGGAAATGTGTTTACGAA CCCTCTTTTACGGGATCAGATTTGAAGCAG	
Qy	1534	GGCGGATTGGTTTTAAATTGCATGAAGACTGGGGSCAACACCAAGTGCATCGATCCACT	1593
Db	1406		1465
		GGGCGATTGGTTTTAAAAATCCACGAAGACTGGGGGAAGCACACCTGCAGCTATTTCACCACT	
Qy	1594	GCATTGACGGTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGATACAGTCAATG	1653
Db	1466		1525
		GCCTCAATGTCGCCGATGAATAGCATGTGCAAGTGGCTATCCACACCGATACCCTTTAAACG	
Qy	1654	AGCAGGTTATGTAGATGACACCTTAAATGCAATGAACGGGGCGGCCATCCATGCCTTACC	1713
Db	1526		1585
		AGGCGGCTGTGTAGAAGACACCTTAGGCGGATTCGCGGGCGCACCATCCATACCTTCC	
Qy	1714	ACATTGAGGAGCGGGTGGAGGACACTCACTGTATGTTATCACCATGGCAGGCGAGTCTCA	1773
Db	1586		1645
		ACACTGAAGGGGCTGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA	
Qy	1774	ATATTCTACCCCTCTCTCCACACCCCACTATTCCTTATACCATTAATPACGGTTGCAGAAC	1833
Db	1646		1705
		ACATTCTACCGGCTCTACTTAAACCGACCATTCCTTTTACCAAAAAACACTGAAGCCGAGC	
Qy	1834	ACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCATCCGGAGGATTTTACAAAT	1893
Db	1706		1765
		ACATGGGACATGTTAATGTGTGTGCCACACATTGGATAAAAGTATCAAGGAAGATGTGCAGT	
Qy	1894	TTTCTCAAAGCCGTATCCGCGCCGGCTCTATCGCGGCTGAAGATGTCTCCATGATATGG	1953
Db	1766		1825
		TTGCGGATTCGAGATTTCGCCGCCCAACTATCGCGCTGAAAGCACTCCATCATGATATGG	
Qy	1954	GTGTGATTCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGGTGCAGGCGAAGTGATTC	2013
Db	1826		1885
		GGATCTTTTCTATCACCCAGCTCCGACTCTCAGGCTATGGACGCGTAGGCGAGGTGATCA	
Qy	2014	CTCGAATCTGGCAGACTGCGGATAGAAATAAAAAAGNAATTTGGTAAAGTCTCTTGAAGATG	2073
Db	1886		1945
		CACGCACTTTGGCAGACAGCAGACAAAAACAAAAAAGAGTTTGGCGCGCTTGAAGAAGAGAA	
Qy	2074	GCAAGATACGATAATTTCCGCATTAAAGCGCTACATCTCCAAATAACACTATCAACCCCG	2133
Db	1946		2005
		AAGCGGATAACGACAACCTCCGCATCAACCGCTACATCTCTAAATAACCATCAACCCCG	
Qy	2134	CTTTTGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCCCGACT	2193
Db	2006		2065
		GGATCGCGCATGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGGCAAAATACGCCGACC	
Qy	2194	TGTTGGTGTGGAAATCTGCGCTTTTGGCGGTAAAAACCAAAATCTGTGATCAAAAGGCGGTA	2253
Db	2066		2125
		TCGTGCTTTTGGAGTCCGGCTTTCTTTTGGCATTAAGCCCAATATGATTTATTAAGGGGGGAT	
Qy	2254	TGTTGGTCTTCTCTGAAAATGGGGGATTTCTAAACGCTCTGTGCCACTTCCCCCAACCGGTTT	2313
Db	2126		2185
		TTATTGGGCTCTCTCAAAATGGGGGATGCCAATGGCTCTATTCCCAACCCCTACGCCGCTCT	
Qy	2314	ATTACCGCAAAATGTTTGGGCATCACGGCAAGCGGAAATTTGACACGAGCATCACTTTTGG	2373
Db	2186		2245
		ATTACCGTGAATGTTTGGACACCATGGGAAAAACAAATTCGACACCAATATCACTTTTCG	
Qy	2374	TTTCCAAAGTCGCTATGAAATATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAGTTTC	2433
Db	2246		2305
		TGTCCCAAGCGGCTTACAGGCGAGGATCAAGAAAGAACTAGGGCTAGATTCGCGCGGCAC	
Qy	2434	TACGGTCAAAAACTCGCGTAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAACCG	2493
Db	2306		2365
		CGCCACTGAAAAACTGTGCGCAATATCATATAAAAGGACCTCAAAATTCACAGATGTGACCG	
Qy	2494	CAAAAAATCACCGTTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCT	2553

CC	sequence, therefore all encoded peptides are cross-referenced to both sequences.									
XX										
SQ	Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;									
	Query Match	36.7%;	Score	1058.6;	DB	5;	Length	4824;		
	Best Local Similarity	65.7%;	Pred. No.	2.2e-227;						
	Matches 1606;	Conservative	0;	Mismatches	799;	Indels	38;	Gaps	3;	
Qy	197	GAGTAAATAGTGAACCTCACACCAAGAGCAAGAAAGTCTTGTTATATATATGCGGC	256							
Db	77	GAGATCTCCATGAACTCACCCCAAGAGTTAGATAGTTGATGCTCCACTACGTGGA	136							
Qy	257	GAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGGAAGCGATGCT	316							
Db	137	GAATTGGCTAAAAAAGCGAAAGAAAGGCAATTAAGCTTAATATGTAGAAAGCAGTAGCT	196							
Qy	317	TACATAGTGCCTATATTAGGACGAGCGCGCTGGAAAGAAACCCGTTGCCAGCTT	376							
Db	197	TTGATTAGTGCCTATATTAGGAAAGCAGAGCTGGTAAAGAACTGCGGCTGAATTG	256							
Qy	377	ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGGTGGGTAAATAG	436							
Db	257	ATGCAAGAGGGCGCACTCTTTTAAACACAGATGATGTGATGGCTGGCAGCATG	316							
Qy	437	GTTCCCGATCTAGGTGTAGAGCCACCTTTCTGATGGTACGAACTTGTAACTGTGAAT	496							
Db	317	ATCCATGAAGTGGTATTGAAGCGATGTTTCTGATGGGACTAAACTCGTAAACCGTGCAT	376							
Qy	497	TGCCCCATCGAACAGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGGGTAA	556							
Db	377	ACCCCTATTGAGGCCAATGTTAAATTTAGTTCTGGTGAGTTG---TTCTTAAAAAATGAA	433							
Qy	557	GACATCGAGCTCAATGAGCAAGAAAGTAAACGAACTTGAGGTACTAATGAAGGCGCT	616							
Db	434	GACATCACTATCAAGNAGCAAAAGCCGTTAGCGTGAAGTTAAATTTGTGGCGAC	493							
Qy	617	AAATCTCTGATGGGTAGCCATTTCCACTCTTTTGAAGCTAACAGGCCACTAAAAATC	676							
Db	494	AGACCGGTTCAATTCGGCTCACACTTCCATTTCTTTGAAAGTGAATAGATGCTAGACTTT	553							
Qy	677	GATCGTGAAGGCTATGGCAACCGCTAGATATTCCTCTTGGCAACAGCTTAGCATTT	736							
Db	554	GACAGAGAAAAAATTTTCGGTAAACGCTTAGACATTTGCGAGCGGACAGCGGTAAAGTTT	613							
Qy	737	GGGCGAGCAAAACCCGAAAGTGCAGTTGATTCCTCTTGTGTGCACTAAAAAGTGAAT	796							
Db	614	GAGCTGGCGAAGAAAAATCCGTAGAAATGATTGACATTTGCGGTAAACAGAAATCTTT	673							
Qy	797	GGCATGAACGGGCTTGTGAATAACATCGCGGATGAACCGCCATAAACATAAAGCGCTTGAC	856							
Db	674	GGATTTAACGATTTGGTTGATAGACAGACAGACAAACGAAGCAAAAAAATTTGCTTTACAC	733							
Qy	857	AAGCGAAATCTCACGATTT-----ATCAAGT	884							
Db	734	AGAGCTAAAGAGCGTGGTTTTTCATGGCGCTAAAAAGCGATGACAACTATGTAAAAACAAAT	793							
Qy	885	AAGGAGACTCCCATGAAA---ATGAAAAACAAGATATGTAATACCTACGGACCCACC	941							
Db	794	AAGGAGTAGAANAATGAAAAAGATTAGCAGAAAAAATATGTTTCTATGTATGGTCTTACT	853							
Qy	942	AAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTAT	1001							
Db	854	ACAGCGCATAAAGTAGATTGGGCGATACAGACTTGATCGCTGAAGTAGAACATGACTAC	913							
Qy	1002	ACCACTTATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGT	1061							
Db	914	ACCAATTTATGCGGAAGAGCTTAAATTCGTTGGCGGTAAAAACCCCTAAGAGAAGGCATGAGC	973							
Qy	1062	CAGNCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCATCACTAAGCGCATGATATC	1121							
Db	974	CNAATCTAACCAACCTTAGCAAAAGAGAGTGGATTTTAATTTATCTACTAACGCTTTTAACGTG	1033							

Qy	1122	GACTACACCGGATTTTCAAAAGCGGACATTTGGGNTTAAAAACGGCAAAATCCATGGCATT	1181							
Db	1034	GATTTACACCGGTATTTTATAAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCATT	1093							
Qy	1182	GGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGGTGGTGGGTGTG	1241							
Db	1094	GGTAAGGCGGTAAACAAGACATGCAAGATGGCGTTAAAAACAATCTTAGCGTAGGTCT	1153							
Qy	1242	GGCACAAGACCATAGCAGGGGAAGGATGATTATTAACCGCTGGGGAAATCGATTCACAC	1301							
Db	1154	GCTACTGAAGCCTTAGCCGGTGAAGTTTGATGTAACGGCTGGTGTATTGACACACAC	1213							
Qy	1302	ACCCACTTCTTTCTCACAACAATTCCTACCGCTAGCCATGCGGTGTACACCATG	1361							
Db	1214	ATCCACTTTATTTCACCCCAACAAATCCCTACAGCTTTTGAAGCGGTGTAAACCAATG	1273							
Qy	1362	TTTGGAGCGGCGCAGGCTCTGTAGATGGCAAGATGCGACTACTATCACTCCCGGCAAA	1421							
Db	1274	ATTGGTGGTGAACCGGTCTCTGTGATGGCACTAATGGGACTACTATCACTCCAGGCAGA	1333							
Qy	1422	TGGAACCTTGCAACCGCATGTTTGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTG	1481							
Db	1334	AGAAAATTTAAAAATGGAATGCTCAGAGCGCTGAAGAAATATTCTATGAATTTAGGTTCTTG	1393							
Qy	1482	GGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGCGATT	1541							
Db	1394	GCTAAAGGTAAACGCTTCTAAACGATGCGAGCTTAGCCGATCAAAATGGAAGCGGTGCGATT	1453							
Qy	1542	GGTTTTAAATTTGCATGAAGACTGGGGCACAACCAAGTGGATCGATCACTGCTTTGAGC	1601							
Db	1454	GGCTTTGCAATTCAGGAAGACTTGGGGCACCACTCTTCTGCAATCAATCATGCGTTAGAT	1513							
Qy	1602	GTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAAGCAGGT	1661							
Db	1514	GTTGGCGACAAATACGATGTGCAAGTCGCTATGCCACAGACACTTTTGAATGAAGCCGGT	1573							
Qy	1662	TATGTAGATGACACCTTAAATGCAATGAAACCGGGCGGCGCATCCATGCGCTACACATGAG	1721							
Db	1574	TGTGTAGAAGACATAATGGCTGCTATTGCTGGACGCACTATGCACTTTTCCACACTGAA	1633							
Qy	1722	GGAGCGGTGAGGACACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTA	1781							
Db	1634	GGGCTGGCGCGGACACGCTCTCTGATATTAATTAAGTAGCCGTGAACACAACTTTCTT	1693							
Qy	1782	CCCTCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGTTGCGAGAACACTTAGAC	1841							
Db	1694	CCGCTTCCACTAACCCCACTCCCTTTACCGTGAATACAGAAGCAGACACATGAC	1753							
Qy	1842	ATGCTCATGACATGCGACCACTAGACAAACGATCCCGGAGGATTTTACAATTTTCTCAA	1901							
Db	1754	ATGCTTATGGTGTGCCCACTTGGATAAAAAAGCATTTAAAGAAGATGTTTCAGTTGCTGAT	1813							
Qy	1902	AGCGTATCCGCCCGGCTCTATCGCGCTCAAGATGTCTCCATGATATGGGTGTGATC	1961							
Db	1814	TCAAGGATCCGCCCTCAAAACCAATTTGCGGCTGAAGACACTTTGCAATGACATGGGATTTTC	1873							
Qy	1962	GCGATGACAAAGCTCGGATTTGCAAGCAATGGGGGTGCGAGCGAAGTGAATTCCTCGAAT	2021							
Db	1874	TCAATCACCGATTTCTGACTCTCAAGCGATGGCGGTGGGTGAAGTTATCACTAGAAT	1933							
Qy	2022	TGGCAGACTCGGATAGAATAAAAAAATTTGGTAAAGTCTTCCGTAAGATGGGCAAGAT	2081							
Db	1934	TGGCAACAGCTGACAAAAACAAGAAATTTGGCCGCTTGAAGAAGAAAAAGCGCAT	1993							
Qy	2082	AACGATATTTCCGATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACC	2141							
Db	1994	AACGACAACTTCAGGATCAAAACGCTTACTTGTCTAAATACACCAATTAACCCAGCGATCGCT	2053							
Qy	2142	CACGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCCGACTTGTGTGGTG	2201							
Db	2054	CATGGGATTACCGAGTATGTAGGTTTCAAGATGGGCAAGTGGCTGACTTGGTATTG	2113							
Qy	2202	TGGAATCCTGCTTTTTTTTGGCGTAAACCCCAAAATCGTGAATAAAGCGGATATGTGGTTC	2261							

Db 2114 TGGAGTCCAGCATCTTTGGCGTGAACCCCAACATGATCATCAAGCGCGATTTCATGCG 2173
QY 2262 TTCTCTGAATGGCGGATTTAAACGGCTGTGCGCCACTCCCAACCGGTTTATTACCGC 2321
Db 2174 TTAAGCCAATGGCGATGGCAAGCTTTCTATCCCTACCCCAACCGGTTTATTACAGA 2233
QY 2322 GAAATGTTGGGATCACCGCAAGCGAAATTTGACACGAGCATCTTTTGTTCCAA 2381
Db 2234 GAAATGTTGGCTCATCATGTAAGCTAAATACGATGCAACATCACTTTGTGTCTCAA 2293
QY 2382 GTGCGCTATGAATGGCTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGTC 2441
Db 2294 GCGGCTTATGACAAAGGCAATTAAGAAGAAATTAGGACTTGAAGACAGAGTTGCGCGTA 2353
QY 2442 AAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGCAAAATC 2501
Db 2354 AAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGCAAAATC 2413
QY 2502 ACCGTGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC 2561
Db 2414 GAAGTCAATCTGAAACTTACCATGTGTTGCGTGGATGCAAGAAAGTAACTTCTAAACCA 2473
QY 2562 ACTGCGAAGTGCCTCTAGCCAGCGCTACACTTTTCTCTAGG 2604
Db 2474 GCCAATAAAGTGAGCTTGGCGCAACTCTTTAGCAATTTTCTAGG 2516

RESULT 12
ABE55112
ID ABE55112 standard; DNA; 4824 BP.
AC ABE55112;
XX
DT 22-SEP-2005 (first entry)
XX
DE Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
XX
KW Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; Immunostimulant; antiulcer;
KW cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX
OS Helicobacter pylori.
OS Salmonella typhimurium.
OS Escherichia coli.
OS Synthetic.
XX
PN WO200132014-A2.
XX
PD 10-MAY-2001.
XX
PF 01-NOV-2000; 2000WO-US030191.
XX
PR 01-NOV-1999; 99US-00431705.
XX
PA (ORAV-) ORAVAX INC.
XX
PI Kleanthous H, Londono-Arcila P, Freeman D;
XX
DR WPI; 2001-343379/36.
DR P-PSDB; ABE55113, ABE55114, ABE55115, ABE55116, ABE55117, ABE55118.
DR ABE55119, ABE55120, ABE55121, ABE55122, ABE55123, ABE55124, ABE55125,
DR ABE55126, ABE55127, ABE55128, ABE55129, ABE55131, ABE55132, ABE55133,
DR ABE55134, ABE55135, ABE55136, ABE55137, ABE55138, ABE55139, ABE55140,
DR ABE55141.
XX
PT Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX
PS Disclosure; SEQ ID NO 1; 63pp; English.
XX
XX The invention relates to inducing an immune response against Helicobacter

CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated Salmonella vector containing a nucleic acid molecule encoding
CC a Helicobacter antigen, and parentally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an htrA or
CC ntrB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
CC represented by both ABE55112 and ABE55130, the peptides expressed by
CC pHUR3 are shown in the sequence listing to be split between ABE55112 and
CC ABE55130, yet Figure 4 shows all the peptides being expressed by the one
CC sequence, therefore all encoded peptides are cross-referenced to both
CC sequences.
XX
SQ Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;
Query Match 36.7%; Score 1058.6; DB 5; Length 4824;
Best Local Similarity 65.7%; Pred. No. 2.2e-227;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;
QY 197 GAGTAATAGGTGAACATCACCCAAAGAGCAAGAAAGTTCTTGTATTATTATGCGGC 256
Db 77 GAGATCTCCATGAACACTCACCCCAAAAGAGTTAGATAAGTTGATGCTCCCTACGCTGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAGCCATTGCT 316
Db 137 GAATTGGCTTAAAAACGCAAGAAAGGCAATTAAGCTTAACTATGTAGAAGCAGTAGCT 196
QY 317 TACATTAGTGCCCATATTATGAGCAAGCGCGCGTGGAAAGAAACCGTTGCCCGACTT 376
Db 197 TTGATTAGTGGCCCATATTATGAGCAAGCGAGAGCTGTTAAAGACTGCGGCTGAATTG 256
QY 377 ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAACCAAGATGATGTGATGCGTGGCAAGCATG 316
QY 437 GTTCCCGATCTAGGTGTAGAGCCACTTCTCTGATGGTACGAACTTGTAACTGTGAAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTTCTCTGATGGGACTAAACTCGTAACCGTGCAT 376
QY 497 TGGCCCATCGAACCCAGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAA 556
Db 377 ACCCTATTGAGGCCAATGGTAAATTAGTTCCTGGTGAGTTG---TTCCTTAAAGAAATGAA 433
QY 557 GACATCGAGCTCAATGCAAGCAAGAAAGTAACCGAACTTGAGGTTACTAATGAAGGCCT 616
Db 434 GACATCACTATCAACGAAGGCAAAAGGCGTTAGCGTGAAGTGTAAAGTGTGGCGAC 493
QY 617 AAATCCTTCGATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAAATTC 676
Db 494 AGACCGGTTTCAAAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCTAGACTTT 553
QY 677 GATCGTGAAAAAGCCCTATGCGAAACCGCTTAGATATTCCTCTCGCAACACGCTACGCATT 736
Db 554 GACAGAGAAAAAATCTTCGGTAAACCTTAGACATTTGACATTTGGCGGTAAACAGATCTTT 613
QY 737 GGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAAAAGTGAAT 796
Db 614 GAGCTGGCGAAGAAAAATCCGTTAGATTTGATTGACATTTGGCGGTAAACAGAGATCTTT 673
QY 797 GGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGGCATTAACATAAGGCTTGAC 856
Db 674 GGATTTAAGCATTTGGTTGTATAGCAAGCAGACAAAGAAAGCAAAAAATTTGCTTTACAC 733

QY 857 AAGCGAATCTCACGATTT-----ATCAAGT 884
Db 734 AGAGCTAAAGAGCGTGGTTTTCTAAGCGCTAAAGCGATGACAACTATGTGTAATAAAACAATT 793
QY 885 AAGGAGACTCCCATGAAA---ATGAAAAACAAGAATATGTAATAACTACCTACGAGCCACC 941
Db 794 AAGGAGTAAGAANTGAATAAGATTAGCAGAAAGAAATATGTTTCTATGTATGCTCTACT 853
QY 942 AAAGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAAGATAGAACATGACTAT 1001
Db 854 ACAGGCGATAAAGTGAATGGGCGATACAGACTTGATCCTGAAAGTAGAACATGACTAC 913
QY 1002 ACCACTATGGCGAAGAACTTAATTTGGCGGCTAAACTATCCGTGAGGCTATGGCT 1061
Db 914 ACCATTTATGGCGAAGAGCTTAATTTGGGTGGCGTAAACCCTAAGAGGAGCATGAGC 973
QY 1062 CAGAGCAATAGCCCTGATGAAACACACCCTAGATTTAGTCACTCACTAACCGGATGATTATC 1121
Db 974 CAATCTAACACCTAGCAAGAAGAGTTGGATTTAATTTACTAACCGCTTTAATCGTG 1033
QY 1122 GACTACACCGGGATTTACAAAGCGCATTTGGATTTAAACACGGCAAAATCCATGGCAAT 1181
Db 1034 GATTACACCGGATTTATAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCAAT 1093
QY 1182 GSCAAGCGGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTG 1241
Db 1094 GGTAAAGGCGGTAAACAAGACATGCAAGATGGCGTTTAAACAACTCTTAGCTAGGTCT 1153
QY 1242 GGCACAGAGACCTAGCAGGGGAAGGTATGATTTATTACCGCTGGGGAATCGATTACAC 1301
Db 1154 GCTACTGAAGCCTTAGCGGTGAAGGTTTGATCGTAAAGGCTGGTATTTGACACACAC 1213
QY 1302 ACCCACTTCCTTCTCCAAACAATTCCTACCGCTCTAGCCAAATGGCGTTTACAACCAATG 1361
Db 1214 ATCCACTTTTACCCCAACAAATCCCTACAGCTTTTGCAAGCGGTGTAAACCAATG 1273
QY 1362 TTTGGAGCGGACAGGTCTGTAGATGGCAGCAAGATCGGACTACTATCACTCCGGGCAA 1421
Db 1274 ATTTGGTGGTGAACCGGCTCGTGTATGGCACTAAATCGGACTACTATCACTCCAGGCGA 1333
QY 1422 TGGAACTTGACCGCATGTTGCGCGAGCAGAGAGTATTCTATGAATGGGCTTTTG 1481
Db 1334 AGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAATTTCTATGAATTTAGGTTTCTTG 1393
QY 1482 GGCAAGGCAATAGCTTAGCAAAAAACAATTTGTAGACAAAGTAGAAGCGGGCGCGATT 1541
Db 1394 GCTAAGGTAAAGCTTCTAAGCATGCGAGCTTAGCCGATCAATTTGAAGCGGTGCGATT 1453
QY 1542 GGTTTTAAATTCATGAAGACTGGGGCAACAACAAGTGCATCTCATCTGCTTGAGC 1601
Db 1454 GGCTTTGCAATTCAGGAAGACTGGGGCACCACTCTTCTGCAATCAATCATGCGTTAGAT 1513
QY 1602 GTGGCAGATCAANTACGATGTCAGTTTGTATCCACACCGATACAGTCAATGAGCGAGT 1661
Db 1514 GTTGGCAGCAAAATACGATGTCAGTGCCTATGCCACAGACACTTTGAATGAAGCGGT 1573
QY 1662 TATGTAGATCACCCCTTAAATGCAATGAAGCGGCGCCATCCATGCTTACCACATTTAG 1721
Db 1574 TGTGTAGAAGACATATGGTGCTATTGCTGGAGCGACTATGCACATTTCCACACTGAA 1633
QY 1722 GGAAGCGGTGGAGGACACTCACTGTATGTTATCACTATGGCAGCGAGCTCAATTTCTA 1781
Db 1634 GGGCTGGCGGAGACGCTCCTGTATATTAAAGTAGCCGTGAACACAACATTTCT 1693
QY 1782 CCCTCTCCAGCACCCCACTATTCCTTATACCATTTAATAGCGTTGACAGACACTTAGAC 1841
Db 1694 CCGCTTTCCATCAACCCCACTATTCCTTCCCGTGAATACAGAAGCAGAGCATGGAC 1753
QY 1842 ATGCTCATGATGCCACCACTAGACAAAGCGCATCCGAGGATTTTACAAATTTCTCAA 1901
Db 1754 ATGCTTATGTTGCCACCACTTGATTAAGCAATTAAGCAATGTTTACGTCGCTGAT 1813
QY 1902 AGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCAATGATATGGGTGTGATC 1961

Db 1814 TCAGGATTCGCCCTCAACCAATTTGGGCTGAAGACACTTTTGCATGACATGGGATTTTC 1873
QY 1962 GCATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTCGAACT 2021
Db 1874 TCAATCACCAAGTTCTGACTCTCAAGCGATGGGCGTGTGGTGAAGTTATCCTAGAACT 1933
QY 2022 TGGCAGATCGGATTAAGAAATAAAGAAATTTGGTAAAGCTTCTGAAGATGGCAAGAT 2081
Db 1934 TGGCAACACAGCTGACAAAAAACAAGAAAGAAATTTGGCCGCTTGAAGAAAGAAAAAGGCGAT 1993
QY 2082 AACGATAATTTCCGCTTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACC 2141
Db 1994 AACGACAACTTAGGATCAAAACGCTACTTGTCTTAATACACCATTAACCCAGCGATCGCT 2053
QY 2142 CACGGCGTAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGATGCGCGACTTTGGTGGTG 2201
Db 2054 CATGGGATTAGCGAGTATGAGGTTTCAGTAGAAGTGGCAAGTGGCTGACTTTGGTATTTG 2113
QY 2202 TGGAACTCGCTCTTTTGGGTTAAACCCAAATCGTGATCAAGGCGGTATGGTGGTC 2261
Db 2114 TGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAAGGCGGATTCATTGCG 2173
QY 2262 TTCTTGAATGGCGGATTTCTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCGC 2321
Db 2174 TTAAGCCAAATGGCGGATGCGAACGCTTCTATCCCTACCCCAACCGGTTTATTACAGA 2233
QY 2322 GAAATGTTTGGGATCAACGGCAAGGGGAAATTTTGACACCGACATCCTTTTGTTCAAA 2381
Db 2234 GAAATGTTTCGCTCATCATGTTGAAGCTAAATACGATGCAAAACATCCTTTTGTGCTCAA 2293
QY 2382 GTGCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTC 2441
Db 2294 GCGGCTTATGACAAAGCATTAAGAAAGAAATAGGACTTGAAGACAAAGTGTGCGGTA 2353
QY 2442 AAAAAGCTGCGTAACATCAACAGAAAGACTTCAAGTTCAACGACAAACCGCAAAATC 2501
Db 2354 AAAAATGCAAGAAATATCACTAAAGAAAGACATGCAATTCACAGACACTTACCGCTCACATT 2413
QY 2502 ACCGTTCATCCGAAACCTTTTCAGGCTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC 2561
Db 2414 GAAGTCAATCCTGAACTTACCATGTGTTCTGTGATGGCAAGAAAGTAACTTCTTAAACCA 2473
QY 2562 ACCTCGCAAGTGCCTTAGCCAGCGGTACACTTTTCTTAGG 2604
Db 2474 GCCAATAAAGTAGCTTTGGGCAACTCTTTAGCATTTTCTAGG 2516

RESULT 13
AAQ12485
ID AAQ12485 standard; DNA; 2767 BP.
XX AC AAQ12485;
XX AC
XX AC
DT 27-AUG-2003 (revised)
DT 23-SEP-1991 (first entry)
XX AC
XX AC
DE DNA encoding A and B subunits of H. pylori urease.
XX Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; ss.
XX Helicobacter pylori.
FH Key Location/Qualifiers
CDS 64..717
FT /*tag= a
FT /label= subunit A
FT /note= "66 kD"
FT 721..2400
FT /*tag= b
FT /label= subunit B
FT /note= "31 kD"
XX XX

PN WO9109049-A.
XX PD 27-JUN-1991.
XX PF 19-DEC-1989; 89GB-00028625.
XX PR 19-DEC-1989; 89GB-00028625.
XX PA (THRE-) 3I RES EXPL LTD.
XX PI Tabaqchal IS, Clayton CL, Wren BW, Kleanthous H;
XX DR WPI; 1991-208084/28.
XX DR P-PSDB; AAR12515, AAR13550.
XX
PT Oligo:nucleotide(s) specific for *Helicobacter pylori* - used as probes and
PT primers to detect *H. pylori* infection, in diagnosis of gastritis, and
PT duodenal and peptic ulceration.
XX
PS Disclosure; Page 16; 28pp; English.
XX
CC The DNA is a 2.7 kb TaqI fragment encoding the A and B subunits of H.
CC *pylori* (previously *C. pylori*) urease, i.e. the 66 and 31 kD antigens.
CC From the sequence, probes and primers can be designed for the
CC amplification (by PCR) of the gene, to produce a prods. common to all *H.*
CC *pylori* strains so far tested and which do not occur in other ureases and
CC can therefore be used as a specific indication of the presence of *H.*
CC *pylori*. *H. pylori* (previously *C. pylori*) is strongly implicated in the
CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
CC primers/probes can be used for the detection of *H. pylori* DNA in gastric
CC mucosa, saliva or faecal samples to provide an early diagnosis of
CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
CC OS field.)
XX
SQ Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;

Query Match 36.5%; Score 1053.6; DB 2; Length 2767;
Best Local Similarity 65.7%; Pred. No. 2.5e-226;
Matches 1615; Conservative 0; Mismatches 804; Indels 39; Gaps 4;

QY 182 AAAATTAAACAAGAGTAATAGTGAACCTCACCCCAAGAGCAAGAAAGTTCTTG 241
DB 40 ACACCTTAAAGATAGGAGTAATGAGATGAACCTCACCCCAAGAGTTAGACAAAGTTGATG 99
QY 242 TTATATATTGCGGCAAGTGTGAGAGCCGAGAGCGAGGCTTAAGCTCAACCAA 301
DB 100 CTCCTACTGCTGGAGAAATAGCTAAACGCAAGCAAGAAAGAGCAATTAAGCTTAACCTAT 159
QY 302 CCGAAGCCATTTGCTTACATTAAGTGCCTATATTATGACGCAAGCGCGCTGGAAAAAAA 361
DB 160 GTGGAAGCGGTACGTTTGATTAGTGCCTATATTATGGAAGAGCGAGAGCTGTTAAAG 219
QY 362 ACGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATCAAGTAATGCC 421
DB 220 ACTGCGCTGAATTGATGCAAGAGGCGCACTCTTTTAAACCGGATGATGTGATG 279
QY 422 GGGTGGGTAAATATGGTTCCCATCTAGGTGTAGAGCCACCTTTCTCTGATGTCAGAA 481
DB 280 GCGGTGGCAAGCATGATCCATGAAGTGGGTATTTGAAGCGATGTTTCTGATGGACCAA 339
QY 482 CTGTGAACCTGTAATTTGGCCATCGAACCAAGATGACACTTCAAGCGGCGCAAGTGA 541
DB 340 CTCGTNACCGTGCATACCCCTATTGAGGCCAATGGTAAATAGTTCTCTGTGATGTTG 396
QY 542 TTTGGTTGCGATAAAGACATCGAGCTCAATGAGCGCAAGAAAGTAAACCGAATCTGAGGTT 601
DB 397 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAGCCGTTAGCGGTGAAGTT 456
QY 602 ACTAATGAAGGCGCTAAATCTTTGATGTGGGTAGCCATTTCCATCTTTTGAAGCTAAC 661
DB 457 CCCCCTGTTGGGACAGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 516
QY 662 AAGGCACATAAATTCGATCGTGAAGAGCCTATGGCAACGCGCTAGATATTCCCTCTGGC 721

DB 517 AGATGCTTAGACTTTGACAGAGAAAAAACTTTCCGTAAACGCTTAGACATTTGCGAGCGG 576
QY 722 AACACGCTACGCTATGGGCGAGGACAAACCCGAAAGTCAGATTGATTCCTCTTTGGTGGC 781
DB 577 ACAGCGGTAAAGTTTGAGCCTCGGAAGAAAAATCCGTAGATTGATGCAATTTGGCGGT 636
QY 782 AGTAAAAAGTGAATGGCATGAACCGGCTTTGTGAATAACATCGCGGATGAACGCCATAAA 841
DB 637 AACAGAAAGTAATCTTTGGATTAAACGCGTTGTTGATAGGCAAGCCGATTAACGAAGCAA 696
QY 842 CATAAAGCGCTTGACAAGCGGAATCTCACGGATTT----- 877
DB 697 AAAATTGCTTTACACAGAGCTAAAGAGCGTGGTTCATCGCGCTAAAGCGATGACAAC 756
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA-----ATGAAAAAACAAGATAATGTAAAT 926
DB 757 TATGTAAACAATTAAGGATGAAGAAATGAAGAAATAGCAGAAAAAGATATGCTTCT 816
QY 927 ACCTACGGACCCACCAAGSCGATAAAGTCGCTTAGGAGATACCGATCTTTTGGGCGAAA 986
DB 817 ATGTATGGCCCTACTACAGSCGATAAAGTGAGATTGGCGGATACAGACTTGTATCGCTGAA 876
QY 987 GTAGAACATGACTATACCACTATGGCAAGAACTTAAATTTGGCGGGGTAAACCTATC 1046
DB 877 GTAGAACATGACTATACCACTTATATGGTGAAGAGCTTAAATTTGGCGGCGGTAAACCCCTA 936
QY 1047 CGTGAAGGTATGGTCAAGAGCAATAGCCCTGATGAACAAACCCCTAGATTTAGTTCATCACT 1106
DB 937 AGAAGAGCAGTACCACTTAAACCCCTAGCAAGAAAGAACTGGATCTTATCATCACT 996
QY 1107 AACCGCATGATTTACGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAACCGGC 1166
DB 997 AAGCTTTAATCGTGGATTACACCGGTAATTTATAAGCGGATATTGTTATTAAGATGGC 1056
QY 1167 AAAATCCATGGCATTTGGCAAGGAGGAAACAAGACATGCAAGATGGCGTTAAGCCCTCAT 1226
DB 1057 AAAATCGCTGGCATTTGTTAAAGCGGTAAACAAAGACACGCAAGATGGCGTTTAAACAAAT 1116
QY 1227 ATGGTCTGGTGTGGGCGAGAGCACTAGCAGGGGAGAGTATGATTATTTACCGCTGGG 1286
DB 1117 CTTAGCGTGGGTCTCTGCTACTGAAAGCCTTAGCCGGTGAAGGTTGATTTGAATCTGCTGGT 1176
QY 1287 GGAATCGATTACACACCACTTTCTTTCTCCACAACAAATTCCTACCGCTCTAGCCAAAT 1346
DB 1177 GGTATTGACACACATCCACTTCTCTCCCCCAACAAATCCCTACAGCTTTTGAAGC 1236
QY 1347 GCGTTTAAACCACTGTTTGGAGCGGCAAGGTCCTGTAGATGGCAAGATGGCACTACT 1406
DB 1237 GGTGTAAACAACCATGATTGGTGGCGAACTGGCCCTGCTGATGGCACTAACGCAACCACT 1296
QY 1407 ATCACTCCGGGCAATTTGGAACCTTGCAACCGATGTTGGCGCAGCAGAGAGTATTTCTATG 1466
DB 1297 ATCACTCCAGGTAGAGAAATTTAAAAATTCATCTCAGAGCGGTGAAGATATTTCTATG 1356
QY 1467 AATGTGCGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAAACAACCTTTGTAGAACAAAGTA 1526
DB 1357 AACTTTGGTTTCTTGCTTAAAGTAAAGCTTCTTACGATCGAAGCTTAGCCGATCAATTT 1416
QY 1527 GAAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGGCAACCAACCAAGTGGCATC 1586
DB 1417 GAAGCTGTGGATTTGGCTTAAATCCCAAGAGACTGGGGGCAACCACTCTCTTCTGCAATC 1476
QY 1587 GATCACTGCTTGAGCGGAGATGAATAGCATGTGGAAGTTGTATGTCACACCGATACA 1646
DB 1477 AATCATGCGTTAGATGTTGGGCAAAATACGATGTGCAAGTCGCTATTTCCACACAGACAT 1536
QY 1647 GTCAATGAGCGAGTTTATGTAGATGACCCCTAAATGCAATGAACGGCGCGCCATCCAT 1706
DB 1537 TTGAATGAAGCGGTTTGGTGGGAAGACACTATGGCAGCTTATGGCGAGCAGCTATGAC 1596
QY 1707 GCCTACCAATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766

Db 1597 ACTTACCACATGAAGCGGTGCGCGGGACACGCTCCTGATATATTAAAGTGCGCGGT 1656
Qy 1767 GAGCTCAATATTCTACCTCCTCTCCACCACCCCACTATTCCCTATACCAATTAATACGGTT 1826
Db 1657 GAACAACATCTCTCCCGCTTCCACTAACCCTACTATCCTTTACCGTGAATACAGAA 1716
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGCGCACACCTAGACAAACGCTCCCGAGGAT 1886
Db 1717 GCGAACAACATGACATGCTTATGCTGTGCACCACTTGATATAAGCATTAAGAAGAT 1776
Qy 1887 TTACAAATTTTCTCAAGCGGTATCCGCGCGCTCTATCGCGCTGAAGATGTCCTCAT 1946
Db 1777 GTCAGTTCGCTGATTCAGAGATTCCCTTCAAAACCAATTCGCGCTGAAGACAATTCGAT 1836
Qy 1947 GATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGGCGAA 2006
Db 1837 GACATGGGATTTTCTCAATCACTAGTTCGACTCTCAAGCGATGGCGTGTGGGTGAA 1896
Qy 2007 GTGATTCCTGAACTTGGCGAGACTGGGGATAAGAAATAAAAGAAATTTGGTAAGTTCCT 2066
Db 1897 GTTATCACTAGAACTTGGCAACAGCTGACAAAAATAAAAAAGAAATTTGGCCGCTTGAA 1956
Qy 2067 GAGATGGCAAGACATACGATATTTCCGCAATTAAGCGCTACATCTCCAAATACATATC 2126
Db 1957 GAAGAAAAAGCGGATACGCAACTTCAGGATCAAAAGCTACTTGTCTAAATACACATT 2016
Qy 2127 AACCCCGCTTTGACCAACGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATC 2186
Db 2017 AACCCAGCATCGCTCATGGATTAGCGAGTATGTCGGTTCTGTAGAAGTGGGCAAGTG 2076
Qy 2187 GCCGACTTGGTGTGGGAATCTGCTTTTGGGTGAATAAAACCCAAATCGTGATCAAA 2246
Db 2077 GCTGACTTGGTATTTGGAGTCCCGCAATCTTTGGTGTGAACCCCAACATGATCATCAA 2136
Qy 2247 GCGGTATGCTGCTTCTCTGAATGGGCGATTTCTTAACGGCTGTGCCACTCCCCAA 2306
Db 2137 GCGGGGTTCATCGCATTTGAGTCAAAATGGGTGTCGGAACGCTTCTATCCCTACCCCAA 2196
Qy 2307 CCGGTTTATPACCGGAAATGTTTGGGCATCACGGCAAGCGGAAATTTGACACAGATC 2366
Db 2197 CAGTTTATACAGAGAAATGTTGCTCATCATGTTAAAGCTAAATAGATGCAACATC 2256
Qy 2367 ACTTTGTTTCCAAAGTCGCTATGAAAATGGCGTGAAGAAAAGCTGGCTTAGAGCGC 2426
Db 2257 ACTTTGTTGCTCAAGCGCTTATGACAAAGGCATTAAAGAAAGATTAGGCTTGAAGA 2316
Qy 2427 CAAGTTCTACGGTCAAAAACCTGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2317 CAAGTGTTCGCGGTAAAAAAATGCGAAATATCACTAAAAAAGACATGCAATTCACGAC 2376
Qy 2487 AAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCCTTTGTAGATGGCAAACTC 2546
Db 2377 ACTACCGCTCACAATTGAAGTCAATCTTGAAACTTACCATTGTTCGTGGATGGCAAA-GA 2435
Qy 2547 TGCACCTCTAAACCCCACTCGCAAGTCCTCTAGCCCAAGCGCTACACTTTCTCTAGG 2604
Db 2436 AGTAACCTTAACACCGCAATAAAGTAGCTGCTGGCGCAACTCTTTAGCAITTTCTAGS 2493

RESULT 14

ADQ37850
ID ADQ37850 standard; DNA; 1710 BP.
XX AC ADQ37850;
XX DT 07-OCT-2004 (first entry)
XX DE H. bizzozeronii ureB DNA.
XX KW Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX

OS Helicobacter bizzozeronii.
XX Key Location/Qualifiers
FT CDS 1. .1710
FT /*tag= a
FT /product= "H. bizzozeronii ureB"
XX US2004142343-A1.
PN 22-JUL-2004.
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHUJ/) ZHU J.
XX Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
DR P-PSDB; ADQ37851.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
PS Claim 7; SEQ ID NO 4; 40pp; English.
XX The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents H. bizzozeronii ureB DNA.
XX SQ Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;
Query Match 32.6%; Score 940.8; DB 12; Length 1710;
Best Local Similarity 72.0%; Pred. No. 4.8e-201; Indels 0; Gaps 0;
Matches 1227; Conservative 0; Mismatches 477;
Qy 900 AAAATGAAAAACAAAGAAATATGTAATACCTACGAGCCCAAGGCGGATAAAGTCGCG 959
Db 7 AAATCTCTCGAAAAGAAATATGTTCTATGATGAGCCCACTACGGGCGATAAAGTGAGA 66
Qy 960 TTAGAGATACCGATCTTTGGGCGAAGATGAGAAATGATATACCATCTATGGCGAAGAA 1019
Db 67 TTGGCGATACCGACCTGATCTTGAAGATCGCAATGACTGCACCACTTATGGCGAAGAA 126
Qy 1020 CTTAAATTTGCGCGGTAAACATATCGTCAGGGTATGGGTCTGAGCAATAGCCCTGAT 1079
Db 127 ATTAAGTTTGGTGGCGGTAAACCACTTCGCGATGGGATGGCAACCAACAGCCCCAGC 186
Qy 1080 GAAACACCCCTAGATTTTAGTCATCACTAAACCGGATGATTATTCGACTACACCGGGATTTAC 1139
Db 187 AGCCACGAACTCGATCTTGCTCACTAAACGCCCTGATCGTGGATTACACCGGCATTAT 246
Qy 1140 AAAGCCGACATTTGGGATTTAAAAACCGCAAAATTCATGTCATTTGGCAAGCAGCAACAAG 1199

Db 247 AAAGCCGATATTGGCATTAAAAATGGCAAAATCATGGCAATTGGCAAAAGCAGGCAATAAA 306
Qy 1200 GACATGAAGATGCGTAAGCCCTCATATGTCGTGGGTGTGGGCACAGAAAGCACTAGCA 1259
Db 307 GACATGAAGATGCGTTTGCAAAATCTTTGGGTGGGCCCTGCTACTGAGGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATGATTCAACACCCACTTCTTCTTCCA 1319
Db 367 GCTGAAGGGCTGATTGTTACAGCTGGTGGGATTGACACCCACATCCACTTTATTTCCTCC 426
Qy 1320 CAACAAATCCCTACCGCTCTAGCCAATGGCGTTTACACCATGTTTGGAGGCGGCACAGGT 1379
Db 427 CAACAAATCCCAAGCATTTGGCAGCGGATCAACACCATGATTGTTGGGGGAACAGGT 486
Qy 1380 CCTGTAGATGGCAAGTACGACTACTATCACTCCGGGCAAAATGGAACTTGGCACCGCATG 1439
Db 487 CCAGCTGATGGACTAACGCGACTACCATCACTCCGGGGCGCTGGAACCTTAAACCATG 546
Qy 1440 TTGCGCGCAGAGAGATATTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATFAGCTCT 1499
Db 547 CTCGGTCCCTCTGAAGAATATGCCATGAACCTTGGGCTATTGGGTAAAGGGAATGTGCT 606
Qy 1500 AGCAAAAAACAATTTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTCATGAA 1559
Db 607 TATGAACCCCTCCCTGGTGGATCACTCGAAGCTGGAGCCATTTGGCTTTAAATCCACGAA 666
Qy 1560 GACTGGGGCAACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGTAGCACACCTCGAGCACTACCATTTGTTGAATGTGGCTGCAAAATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTTATGTPAGATGACCCCTA 1679
Db 727 GTGCAAGTGGCTATCCACCGATACCTTGAATGAAGCGGGCTGTGTGGAAGACACTTTG 786
Qy 1680 AATGCAATGAAGCGCGCGCCATCCATCCCTACACATTCAGGAGCGGGTGGAGACAC 1739
Db 787 CAAGCCATTCCTGGGCGCACTATCCACACTTCCACACTGAAGGTGCTGTGGGGGCAC 846
Qy 1740 TCACCTGATGTTATCACATGGCAGCGAGCTCAATATTCTACCTCTCCACACCCCC 1799
Db 847 GCTCCGATGCTAATGAGTGTCTGGCGAATTTAAACATCTCCAGCTTCTACCAACCC 906
Qy 1800 ACTATTCCCTATACCAATTAATACGGTTTGCAAGACACTTACATGCTCATGACATGCCAC 1859
Db 907 ACCATTCTTTCCCGTGATACAGAGCCGACACATGACATGTTGATGTTGCCAC 966
Qy 1860 CACTAGACAAACGATCCCGAGGATTTACAATTTCTCAAGCCGATATCCGCCCGGC 1919
Db 967 CACTTGGATAAAAACATCAAGAAGATGTCAGTTTGTCTGATTTCTAGGATTCGCCCCCA 1026
Qy 1920 TCTATCGCGCTGAGATGCTCCATGATGCTGATGCTGATCGGATGACAGCTCGAT 1979
Db 1027 ACCATCGCGCTGAGGACAAACTCCAGATATGGGATTTTCTATACACAGCTCTGAC 1086
Qy 1980 TCGCAAGCAATGGCGGTGCGAGCGAAGTGAATTCCTCGAACTTTGGCAGACTGGCGATAAG 2039
Db 1087 TCCCAAGCGATGGCCGTGTAGCGAGTCAATCACCGCACTTGGCAACAGCGGCACAA 1146
Qy 2040 AATAAAAAAGAAATTTGGTAGCTTCTGAAAGATGGCAAAAGATACGATAATTTCCGATT 2099
Db 1147 AACAAAAAGAAATTTGGTCTGCTTGGCTGAGGAAATTCGCCGATTTGGTCTCTGCTTTT 1206
Qy 2100 AAGGCTACATCTCAAAATACATCAACCCCGCTTTGACCCACCGCGTGGAGGATAT 2159
Db 1207 AAGGCTACATTTTCAAAATACACCATCAACCCCGCTATTGACACAGCGCATTTCTGAATAT 1266
Qy 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCGCATTTGGTGTGGGAATCTGCTTTTTT 2219
Db 1267 GTGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTCTTTGGAGTCTCTGCTTTT 1326
Qy 2220 GCGTAAAAACCCAAATTCGTGATCAAGGCGGTATGGTCTTCTCTGAAATGGGCGAT 2279
Db 1327 GGCATTAAACCCCAACATGATCATCAAGAGGGGATTCATCGCACTTTCTCTCAATGGGCGAT 1386

Qy 2280 TCTAACCGCTGTGTGCCACTCTCCCAACCGGTTTATTACCCGAAATGTTTGGCATCAC 2339
Db 1387 GCCAATGCTCTATCCCACTCTCCCAACCGGTGTTATTACCGGAAATGTTTGGCCACCAT 1446
Qy 2340 GGCAAGCGAAATTTGACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAAATGGC 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGCAACGGC 1506
Qy 2400 GTCAAGAAAAAGCTGGGCTTTAGAGCGCAAGTTTCTACCGGTCAAAAACTGCCGTAACATC 2459
Db 1507 ATTAAGAGAGTTGGGCTTGCAGAGTGGTTTGGCAGTTAAAACTGCGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAAAATCACCGTCGATCCGAAACC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCACGATGTTTACCGCACATCGAAGTCAATCTCTGAAACC 1626
Qy 2520 TTCGAGTCTTTGTAGATGCAAACTCTGCACCTCTTAAACCCACCTCGAAGTGCCTCTA 2579
Db 1627 TACAAAGTTTAAAGTGGATGGCAAGAGTTTACTTCCAAAGCAGCGATAAAATCAGCCTA 1686
Qy 2580 GCCAGCGCTACACTTTTCTTCTAG 2603
Db 1687 GCACACTCTACAACTTGTTCTAG 1710

RESULT 15

ABR00816

ID ABR00816 standard; cDNA; 1719 BP.

XX ABA00816;

AC AC

XX 01-APR-2003 (first entry)

DT 01-APR-2003 (first entry)

XX H. felis urease B coding sequence.

Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
immune response; circulatory vessel; Gut Associated Lymphoid Tissue;
GALT; immune response; antigen; Salmomella; Cholera; immunity;
Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
hormone; hormone receptor; cancer; ss.

XX Helicobacter felis.

OS Helicobacter felis.

XX WO200296949-A1.

PN 05-DEC-2002.

PD 24-MAY-2002; 2002WO-AU000661.

PF 25-MAY-2001; 2001AU-00005241.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX McKenzie B, Boyle J, Lew A;

XX WPI; 2003-156789/15.

XX Raising an immune response in an animal by administering composition

XX comprising carrier and antigen bound to targeting moiety which binds to

XX receptor present in circulatory vessels in Gut Associated Lymphoid

XX tissue.

XX Disclosure; Page 27-28; 45pp; English.

XX This sequence encodes H. felis urease B. This sequence was used in the

XX method of the invention for raising an immune response in an animal. The

XX method comprises administering to the animal a composition comprising a

XX carrier and an antigen bound to a targeting moiety which binds to at

XX least one receptor present in circulatory vessels in Gut Associated

XX Lymphoid Tissue (GALT). The method is useful for raising an immune

XX response in an animal against antigens from Salmomella, Cholera,

XX Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut

CC associated toxins, gut hormones, gut hormone receptors or gut associated
CC cancers. The method is useful for raising both mucosal and systemic
XX immunity against any antigen used in the composition
SQ Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;

Query Match 31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity 70.6%; Pred. No. 2e-192;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

QY 900 AAAATGAAAAAACAAGATATGTAATACCTACGGACCCACCAAAAGCGATAAAGTGGCG 959
DB 15 AAGATTTCACAAAGAAATATGTTCTATGATATGTTCCCACTACCGGGGATCGTTAGA 74
QY 960 TTAGGAGATACCGATCTTTGGGAGAGTAGAATACATGACTATATACCACTATGGCGAAGAA 1019
DB 75 CTCGGGACACTGATTGATCTTTAGAGTGGAGCATGATTGCACCACTTATGTTGAAGAG 134
QY 1020 CTTAAATTTGGCGGGTAAACTATTCCTGAGGGTATGGTTCAGGCAATAGCCCTGAT 1079
DB 135 ATCAAAATTTGGGGCGGTAAACTATTCCTGATGGGATGATCAAAACCAATAGCCCTAGC 194
QY 1080 GAAAAACCTAGATTTAGTTCATCACTAACCGCATGATTATCGACTACACCGGATTTAC 1139
DB 195 TCTTATGAATTTAGTTGGTGTCTCACTAACGCCCTCAATTTGGACTATACGGGCATTTAC 254
QY 1140 AAAAGCCGACATTTGGGATTAATAAACCGCAAAATCATGGCATTTGGCAAGGAGGAAACAAG 1199
DB 255 AAAGCCGACATTTGGATTAAGACGCGCAAGATTTCAGGCATTTGCAAGGCGAGCAATAAG 314
QY 1200 GACATGCAAGTGGCGTAAGCCCTCATATGGTGTGGGTGGGCAAGAAAGCACTAGCA 1259
DB 315 GACATGCAAGTGGCGTAGATAATAATCTTTGGGTAGGTCTGTCTACAGAGGCTTTGGCA 374
QY 1260 GGGGAGGTATGATTATTACGCTGGGGGATTCAGTTTCAACACCCACTTCTTTCTCCA 1319
DB 375 GCTGAGGCTTGATTGTAACCGGTGGTGGCATCGATACGCAATPTCACTTTATCTCTCCC 434
QY 1320 CAACAAATTCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGTTTGGAGCGGCACAGGT 1379
DB 435 CAACAAATTCCTACTGTTTGGCAGCGGGTTACAACCATGATTGGAGGAGGCACAGGA 494
QY 1380 CTTGTAGATGGCAGAAATCGACTACTATCATCTCCGGGCAAAATGGAACTTGCACCGCATG 1439
DB 495 CTTGCGATGGCAGAAATGGACCACTCATCTCCGACGCGCTAATCTTAAAGATATG 554
QY 1440 TTGGCGCCAGAGAGATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGCTCT 1499
DB 555 TTGGGTGACCGCAAGAAATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAATGTGTCT 614
QY 1500 AGCAAAAAACAACCTTGTAGAACAGTAGAAGCGGGCGGATTTGTTTAAATTCATGAA 1559
DB 615 TACGACCTCTTTACCGGATTCAGATTGAAGCAGGGGATTTGTTTAAATTCACGAA 674
QY 1560 GACTGGGGCAACAACCAAGTGGCATCGATCATCTGTTGAGCGTGGCAGATGAATACGAT 1619
DB 675 GACTGGGGAGCAACACCTGCGAGCTATTTCACCACTGCCTCAATGTCCGCGATGAATACGAT 734
QY 1620 GTCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGATTTATGATGACACCCCTA 1679
DB 735 GTGCAAGTGGCTATTCACACCGATACCCCTTAAACGAGGGGCTGTGTAGAGACACCCCTA 794
QY 1680 AATGCAATGAACGGCGCGCATCCATGCCCTACCAATTGAGGGAGCGGGTGGAGGACAC 1739
DB 795 GAGGCGATTTGGGGGCGGACCATCCATACCTTTCACACTGAAGGGGCTGGGGGTGGACAC 854
QY 1740 TCAGCTGATTTATCACCATGGCAGGAGCTCAATATTTTACCCCTCTCCACCAACCCCC 1799
DB 855 GCTCCAGATTTATCAAAATGGCAGGGGAATTTAACTTTCTACCGCTCTACTTAAACCG 914
QY 1800 ACTATTCCTATACATTAATACGGTTGCAAGAACATTTAGACATGCTCATGACATGCCAC 1859
DB 915 ACCATTCTTTTCAACCAAAAACTGTAAGCCGAGCACATGGACATGTTTAAATGGTGTGCCAC 974

QY 1860 CACTAGACAAACGCAATCCGCGAGGATTTACAATTTTCTCAAAGCCGTATCCGCCCCGGC 1919
DB 975 CACTTGTATAAAGATATCAAGGAAGATGTGCAGTTTGCCTGATTCGAGGATTCGCCCAA 1034
QY 1920 TCTATCGCGCTGAAGATGTCTCCATGATATCGGTGTGATCGGATGACAAGTCCGAT 1979
DB 1035 ACTATCGCGCTGAAGACCAACTCCATGACATGGGATCTTTCTATCACAGTCCGAC 1094
QY 1980 TCCAAAGCAATGGGGGTGCGAGCGAAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAG 2039
DB 1095 TCTCAGCTATGGACGCGTAGCGAGTGTATCACACGCACTTTGGCAGACAGACAAA 1154
QY 2040 AATPAAAAAGAAATTTGGTAACTTCTGAAAGTGGCAAAAGATTAACGATAATTTCCGAT 2099
DB 1155 AACAAAAAGAGTTTGGGCGCTTTGAAAGAGGAAAAAGCGATTAACGACAACTTCGCATC 1214
QY 2100 AAGCGCTACATCTCCAATAACATCAACCCCGCTTTGACCCACGCGTGAGCGAGTAT 2159
DB 1215 AAACGCTACATCTTAATAACACCAATCAACCCCGCATCGGCATGGGATTTCTGACTAT 1274
QY 2160 ATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTTGGTGTGTGGAATCTCTGCTTTTT 2219
DB 1275 GTGGCTCTGTGGAAGTGGCAATACGCGGACTCTGTGCTTTGGAGTCCGGCTTTCTTT 1334
QY 2220 GGCGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGCGAT 2279
DB 1335 GGCATTAAAGCCCAATATGATTATTAAGGGCGGATTTATTGGCTCTCTCAATGGCGAT 1394
QY 2280 TCTAAAGCGCTCTGTGCCACTCCCAACCGGTTTATACCGCAAAATGTTTGGGCATCAC 2339
DB 1395 GCCAATGCGTCTATTCCCACTCCAGCCGCTATTATTACCGTGAATTTTGGACACCAT 1454
QY 2340 GGCAAGGCGAAATTTGACACGACATCACTTTTGTTCCAAAGTCGCTATGAAAAATGGC 2399
DB 1455 GGGAAAAACAATTCGACCAATATCACTTTGTGTCCCAAGCGGCTTACAAGCAGGG 1514
QY 2400 GTGAAGAAAAAGCTGGGCTTTAGAGCGCAAGTCTACCGGTCAAAAACTGCGGTAAACATC 2459
DB 1515 ATCAAGAAAGAACTAGGGCTAGATCGGTGTATTTGCCAGTGAATAAACTGTGCAATATC 1574
QY 2460 ACCAAGAAAGCTTCAAGTTCAAGCAAAAACGGCAAAATCACCGTCGATCCGAAACCC 2519
DB 1575 ACTAAAAAGGACCTCAAAATTCAGCATGTGACCGCACATATTGATGTCAACCCCTGAAAC 1634
QY 2520 TTGAGGTCTTTGTAGATGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
DB 1635 TATAAGGTGAAGTGGATGGCAAGAGGTAACTCTTAAAGCAGCAGATGAATTGAGCCTA 1694
QY 2580 GCCCAGCGCTACACTTTTCTTAG 2603
DB 1695 GCGCACTTTTATAATTTGTTCTAG 1718

Search completed: November 29, 2005, 00:34:36
Job time : 1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:31:00 ; Search time 10210 Seconds
(without alignments)
13211.279 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rrgagatttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gsal.*
10: gb_gsal2.*
11: gb_gsal3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327.4	11.4	958	8	DR637438
2	324.2	11.2	946	8	DR637093
3	310.2	10.8	1509	4	CN50A8R8
4	309	10.7	872	7	CO122599 GR_Eb04F
5	306.2	10.6	925	8	DR633422
6	302.2	10.5	748	8	CX676965 ycd29h04.
7	299.2	10.4	889	8	DR637386
8	296.4	10.3	796	8	DR633939
9	292	10.1	849	8	DR635599
10	279.6	9.7	766	7	CV195867
11	274.2	9.5	986	6	CF826586
12	273	9.5	784	9	AZ935182
13	262.6	9.1	990	7	CO027797
14	261.2	9.1	684	7	CF714848
15	261.2	9.1	921	6	CN907928
16	261	9.1	946	8	DN809412
17	259.6	9.0	713	7	CK757406
18	258.2	9.0	689	8	DR440006
19	255.2	8.9	781	8	DR916924
20	254.8	8.8	863	6	CD375065
21	254.4	8.8	945	8	DN808832
22	249.4	8.7	757	5	BU027777

ALIGNMENTS

RESULT 1
DR637438

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DR637438 958 bp mRNA linear EST 11-JUL-2005
EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
sequence.
DR637438.1 GI:70712272
EST.
Gibberella moniliformis
Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 958)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR

USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMAW64TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES
source

Location/Qualifiers
1..958
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAW64"
/cissue_type="mycelia"
/clone_lib="FvM"

/notes="Vector: pBluescript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was

CA264652 SCQGLB204
CN908212 030109ABL
CF708414 CCABR70TR
CA236757 SCEQL505
CV902222 PD019F1 m
AG940016 Drosophil
DN810553 76954039
BG887260 EST13111
CF689903 CCADS40TF
CN910029 030124ABL
CA851837 D18B02.D1
CX557362 yda35h1.
AW618902 EST120886
CF690511 CCRAZ91TR
AZ935008 BJ_Ba000
BZ572656 msH2 2738
CF706674 CCAF8387O
AZ935371 BJ_Ba000
AG351733 131F03 Ma
AW618900 EST120886
AG939430 Drosophil
CF702527 CCAHQ83TF
AG941093 Drosophil

23 248.2 8.6 905 6 CA264652
24 247.8 8.6 661 7 CN908212
25 247.4 8.6 796 6 CF708414
26 240.8 8.4 789 6 CA236757
27 240.4 8.3 692 8 CV902222
28 239 8.3 643 10 AG940016
29 238.8 8.3 937 8 DN810553
30 237.2 8.2 634 2 BG887260
31 235.4 8.2 829 6 CF689903
32 234.2 8.1 626 7 CN910029
33 233.8 8.1 636 6 CA851837
34 226.2 7.8 708 8 CX557362
35 226 7.8 589 1 AW618902
36 225.2 7.8 747 6 CF690511
37 224.4 7.8 660 9 AZ935008
38 223.8 7.8 1275 9 BZ572656
39 222.4 7.7 890 6 CF706674
40 220.8 7.7 614 9 AZ935371
41 220 7.6 750 2 BG351733
42 219.2 7.6 555 1 AW618900
43 218 7.6 663 10 AG939430
44 218 7.6 878 6 CF702527
45 216.8 7.5 621 10 AG941093

ORIGIN	Query Match Best Local Similarity Matches 565; Conservative 0; Mismatches 376; Indels 2; Gaps 1;	11.4%; Score 327.4; DB 8; Length 958; 59.7%; Pred. No. 3.5e-70; 0; Mismatches 376; Indels 2; Gaps 1;	directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."
QY	1209	GATGGGTAAGCCCTCATATGTCGTGGGTGGGACAGACAGACACTAGCAGGGGAAGGT	1268
DB	2	GACGGGTAACAGAGGCGATGTCGTGCGAAGCTGCACGGATGTCGTAGCAGGTGAAGGA	61
QY	1269	ATGATTATATCCCTGGGGAATCGATTACACACCACTTCCTTTCTCCACAAATTC	1328
DB	62	AAGATTGTGACCGGGCGCTATCGACACCATATTCATTTATCTGCCCCACGACAGTA	121
QY	1329	CCTACCGCTCTAGCCATGCGTTACACACCATGTTGGAGCGGCACAGGTCTCTGTAGAT	1388
DB	122	CCGAAAGCTCTTGATCTGTGTAAACACCATGCTTGGCGGTGTACCGGCCCAAGTGCT	181
QY	1389	GGCAGCAATCGACTACTACTCTCGGGGCAAAATGGAACCTTGACCGCATGTTGCGGCA	1448
DB	182	GGAACGAACCACTACTTGTACGCTGTGCTCATTTACATGCTCAAAATGTTGAGGCG	241
QY	1449	GCAGAGAGTATTCATGATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAA	1508
DB	242	TGGCATCAGCTTCTATCAATATGTTATTTACTGCGCAAGGTAATGATAGTTCTCTCGAG	301
QY	1509	CAACTTGTAGAACAGTAGAGCGGCGGATGTTGTTTAAATTTGCAATGAGAGCTGGGC	1568
DB	302	GGTCTGGCGCATCAGGTCAATGCTGTGCTGTGGCTCAAGCTTCAATGAGGACTGGGGT	361
QY	1569	ACAACACCAAGTGCATCACTGCTTGAGGCTGGCAGATGAATACGATGTGCAAGTT	1628
DB	362	TGCACTCTCTGCTATGAGCTTGTCTCAGTGTCTGTGATGATTCGATATCAATGT	421
QY	1629	TGATTCACACCGATCAGTCAATAGGACGAGTTATGATGATGACACCCTAATATGCAATG	1688
DB	422	CTTATTACACTGACACGCTTAAAGAGTCTGGCTTTTGTGCAATCTACGATCGCTGCTTC	481
QY	1689	AACGGGGCGCCATCCATGCTACCACTTACGAGGAGCGGTGGAGGACACTACCTGAT	1748
DB	482	AAGAACCGCAATATCACTTATCAACAGAGGAGTGCAGGAGGTGCGCATGCTCCGGAT	541
QY	1749	GTTATCACTAGGCGAGGCTCAATATTTCTACCTCTCCACACCCCACTATTCCC	1808
DB	542	ATCATCTCGGTGATGAGCATCAAAATGTTCTGCCATCTGACCAACCCCTACAGACCA	601
QY	1809	TATACCAATTAATAGGTTGAGAACACTTACACATGCTCATGATCGCCACCACTAGAC	1868
DB	602	TTACACGCAATCTCTCGATGAGCATCTCGATATGCTTATGTTCTGCCATCACTTGTCC	661
QY	1869	AAAGCATCCGCGAGGATTTACAAATTTCTCAAAGCGTATCCGCCCGGCTCTATCGCG	1928
DB	662	AAGAATATCCAGAGGATGAGCCCTCGCGAGAGCGGTATTCGTGCTGAAACCAATGCT	721
QY	1929	GCTGAAGATGCTTCCATGATATGAGTGTGATCCGATGATCAAGCTCGATTCGCAAGCA	1988
DB	722	GCTGAGGATGATTACACGCAAAAGGCGCTATACGATGATGAGCTCTGATCTCGAGCT	781
QY	1989	ATGGGGGCTGCAGCGAAGTATTCCTCGAACTTTGGCAGACTGCGGATGAAGATAAAAAA	2048
DB	782	ATGGGCGTTTGCAGAGGTCGTTTAAAGACATGGAATCTGCGCATAAATAAAGTG	841
QY	2049	GAATTTGGTAAAGTCTCTGAGATG - GCAAAGATAACGATAATTTCCGCAATTAAGCGCT	2106
DB	842	CAGAGGGGTTGGTTCGCGAGGATGAAGGACAGGGGCTGATAATGCGCGTGTAAAGCT	901
QY	2107	ACATCTCCAATACATCATCAACCCGCTTTGACCCACGGCT 2149	
DB	902	ATGTCAGCAAGTATACATTAACCCAGCTATTGCTCAGGGCTT 944	

RESULT 2 LOCUS DEFINITION	DR637093 EST1027718 FVM Gibberella moniliformis cDNA clone FVMAT95, mRNA	946 bp moniliformis cDNA clone FVMAT95, mRNA	linear EST 11-JUL-2005
ACCESSION	DR637093	GI:70711927	
VERSION	DR637093.1		
KEYWORDS	EST.		
SOURCE	Gibberella moniliformis		
ORGANISM	Gibberella moniliformis		
REFERENCE	Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.		
AUTHORS	1 (bases 1 to 946) Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D. and Whitelaw, C.A.		
TITLE	Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster		
JOURNAL	Unpublished (2005)		
COMMENT	Contact: Brown, D.W. USDA/ARS/NCAR		
FEATURES	source 1..946 /organism="Gibberella moniliformis" /mol_type="mRNA" /strain="m3125" /db_xref="taxon:117187" /clone="FVMAT95" /tissue_type="mycelia" /clone_lib="FVM" /note="vector: pBlueScript II SK(+) XR; Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FVM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GVAM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."		
ORIGIN	Query Match Best Local Similarity Matches 563; Conservative 0; Mismatches 378; Indels 2; Gaps 1;	11.2%; Score 324.2; DB 8; Length 946; 59.7%; Pred. No. 2.2e-69;	
QY	1209	GATGGGTAAGCCCTCATATGTCGTGGGTGGGACAGACAGACACTAGCAGGGGAAGGT	1268
DB	2	GACGGGTAACAGAGGCGATGTCGTGCGAAGCTGCACGGATGTCGTAGCAGGTGAAGGA	61
QY	1269	ATGATTATATCCCTGGGGAATCGATTACACACCACTTCCTTTCTCCACAAATTC	1328
DB	62	AAGATTGTGACCGGGCGCTATCGACACCATATTCATTTATCTGCCCTCAGCAAGTA	121
QY	1329	CCTACCGCTCTAGCCATGCGTTACACCACTGTTGGAGCGGCACAGGTCTCTGTAGAT	1388
DB	122	CCGAAAGCTCTTGATCTGTGTAAACACCATGCTTGGCGGTGTACCGGCCCAAGTGCT	181
QY	1389	GGCAGCAATGCGACTACTATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGGCA	1448
DB	182	GGAACGAACCACTACTTGTACGCTGTGCTCATTTACATGCTCAAAATGTTGAGGCG	241
QY	1449	GCAGAGAGTATTCATGATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAA	1508
DB	242	TGGCATCAGCTTCTATCAATATGTTATTTACTGCGCAAGGTAATGATAGTTCTCTCGAG	301

QY	1509	CAACTTGTAGAACAGTAGAGCGGCGCATGTTGGTTTAAATTGCATGAAGACTGGGGC	1568
Db	302	GGTCTGCGCGATCAGGTCAATGCTGGTGTGTGGCTCAAGCTTTCAGGAGCTGGGT	361
QY	1569	ACAACACCAAGTGCATCGATCACTGCTTGAGCGTGGCAGATGAATACAGATGCAAGTT	1628
Db	362	TGCATCTCCCTGCTATTGACGGTTGCTCAGTGTCTGTGATGAATTCGATTAATGT	421
QY	1629	TGTATCCACACCGCATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGCAATG	1688
Db	422	CTTATTTCACATGACACGGTTAAACGAGTCTGGCTTTGTGCAATCTACGATCGCTGCTTC	481
QY	1689	AAGCGGCGGCCATCCATGCTTACCATGATGAGGAGCGGGTGGAGGACATCACTCATGT	1748
Db	482	AAGAACCGCACATTCATCTTATCACACAGAGGGTGCAGAGGTGGCCATGCTCCGGAT	541
QY	1749	GTTATCACCATGCGAGCGAGTCAATATTTCTACCTCTCCACACACCCCTATTTC	1808
Db	542	ATCATCTCCGTGTAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTACAAGACCA	601
QY	1809	TATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCACCACTTAGAC	1868
Db	602	TTTCACACGCAATCTCTGATGAGCATCTCGATATGCTTATGGTCTGCCATCACTTGTC	661
QY	1869	AAACGCATCCGCGAGGATTTACAATTTCTCAAAGCGTATCGCCCCGGCTCTATCGCG	1928
Db	662	AAGAATATCCAGAGGATGTAGCTTTCGGGAGAGCGGTATTCTGCTGAAACCAATTGCT	721
QY	1929	GCTGAAGATGCTCCATGATATGGTGTGATCGGATGACAGCTCGGATTCGCAAGCA	1988
Db	722	GCTGAGGATGTATTACACCAAGCGCTATCAGCATGATGAGCTCTGACTCGCAGGCT	781
QY	1989	ATGGGCGTGCAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGGATGAAGATAAAAA	2048
Db	782	ATGGGCGGTGGGAGAGTGTGTTTAAAGACATGGAATCTCGGCATTAATAAATGATG	841
QY	2049	GAAATTTGGTGAAGTCTCTGAAGAT--GGCAAGATAACGATAATTTCCGCATTAAGCGCT	2106
Db	842	CAGAGGGTGGTTCGGGAGGATGAGGCACAGGGGCTGATAATGCGGCTGTAACGCT	901
QY	2107	ACATCTCAAATACACTATCAACCCCGCTTTGACCCACCGCGT	2149
Db	902	ATGTCAGCAAGTATACTATTACCCAGCTATTGCTCAGGGCTT	944
RESULT 3			
CNSOABR8			
LOCUS			
DEFINITION			
CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004			
Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
GSLTPGH4ZD08 of Hormone Treated Callus of strain col-0 of			
Arabidopsis thaliana (thale cress).			
BX816461			
BX816461.1 GI:42471655			
HTC; GSLT cDNA.			
Arabidopsis thaliana (thale cress)			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;			
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
1 (bases 1 to 1509)			
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,			
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,			
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.			
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:			
A Combined Approach to Evaluate and Improve Arabidopsis Genome			
Annotation			
Unpublished			
2 (bases 1 to 1509)			
Genoscope.			
Direct Submission			
Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage :			
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
- Web : www.genoscope.cns.fr)			

COMMENT		The sequences are based on single pass reads.	
		Life Technologies (a division of Invitrogen) members carried out	
		full-length libraries construction : Temple G.	
		Genoscope members carried out sequencing and annotation : Castelli	
		V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,	
		Schachter V., Weissenbach J., Salanoubat M.	
		URGV INRA : Clepet C., Caboche M.	
		Annotation is based on the June 2003 version of the Arabidopsis	
		genome released by MIPS (Munich Information center for Protein	
		Sequences). 5 prime and 3 prime are assembled with Phrap.	
		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full	
		length	
		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.	
FEATURES		Location/Qualifiers	
source		1..1509	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clone="GSLTPGH4ZD08"	
		/tissue_type="Hormone Treated Callus"	
		/ecotype="Col-0"	
		/plasmid="pCMVSPORT_6"	
gene		1..1509	
		/gene="Atlg67550"	
ORIGIN		Query Match 10.8%; Score 310.2; DB 4; Length 1509;	
		Best Local Similarity 59.1%; Pred. No. 7.7e-66;	
		Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;	
QY	1587	GATCCTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATACA	1646
Db	1	GACAATTTGTTGGCAGTTGCAGAAGAATATCATATCCAGGTGAACATTCTACTGACACC	60
QY	1647	GTCATGAGCGCAGTTATGTAGATGACACCTTAATGCAATGAACGGCGCCATCCAT	1706
Db	61	TTAAATGAATCCGGTTTGTGTGAGCACACTATCAATGCAATTCGTGGGAGAACAAATACAT	120
QY	1707	GCTACACCATTCGAGGAGCGGTGGAGGACACTCACCTCATGTTATCACCATGGCAGGC	1766
Db	121	ACATATCAGATTAAGGTGCTGTGGTGGACATGACACAGATATCATTAGAGTTTGTGA	180
QY	1767	GAGCTCAATATTCTACCTCTCTCCACACCCCACTATTCCCTATACATTAAATACGGTT	1826
Db	181	GTGAAAAATGTACTCCCGTCATCAACCAACCCCAACCGCCATATACGAAAAATCTGTA	240
QY	1827	GCAGAACCTTAGACATGCTATGACATGCCACCACTAGACAAACGATCCGCGAGAT	1886
Db	241	GATGAACATCTTGACATGCTGATGGTTGGCCATCACCTTGACAAGAACATCCCAAGAGAT	300
QY	1887	TTACAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCCGGCTGA-AGATGTGCTCCA	1945
Db	301	GTAGCTTTTGTGTAATCAAGAATAAGAGCTGAACAATAGCTGCAGAGAGATATATTGCA	360
QY	1946	TGATATGGTGTGATCGCATCAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGA	2005
Db	361	TGATATGGGAGCAATCAGCATTTATCTCTCCGACTCACAAGCAATGGTCCGATTTGGAGA	420
QY	2006	AGTGATCTCTCGAATCTGGCAGACTGGGGATAGATAAAAAAGAAATTTGGTAGCTTCC	2065
Db	421	GGTAATAAGCAGAACTTGGCAGACAGCTGCAAGATGAAGGCTCAGAGGGGAGCCATTAA	480
QY	2066	TGAAGATGCAAGATGAAGTAATTTCCGATTAAAGC-GCTACATCTCCAAATFACACTA	2124
Db	481	TCTTAACATGGCTGATGATTAACCAAGGATCAAGCATATCATTTGCAAGTTACTACTA	540
QY	2125	TCAACCCCGCTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGTGAAGAGGCGAAGA	2184
Db	541	TAAACCCAGCTATAGCCAATGGGTTTGCACACTTAATTGGCTCGTTGAGGTGAAGAAGC	600
QY	2185	TGCGCGACTTGGGTGTGAAATCCCTGCTTTTTTTT-GGCGTAAACCCCAAAATCGTGATC	2243
Db	601	TGCTGATCTTGTATTATATGCGACCGCGCTTCTTCTCGGAGCAAAACCCAGAAATGATAATC	660


```

QY 2244 AAAGCGGTATGGTGTCTTCTCTGAAATGGCGGATTTAAACCGTCTGTGCCACTCCC 2303
D 661 AAAGGAGGCAATATAGCATGGCTAATATGGCGATGCAAAATCGAGCATACCCACTCCT 720
QY 2304 CAACCGGTTTATACCGGCAATGTTTGGCATACGCGAAGCGGAAATTTGACACCAGC 2363
D 721 GAGCGGTGATATCAAGACCTATGTTTGGAGCATTTGGAAAGGCGGAAAGTGAATACTCT 780
QY 2364 ATCACTTTTGTTCACAAAGTCGCTTATGAAATGGCGTGAAGAAAGCTGGCTTAGAG 2423
D 781 GTTGCAATTTGTCAGCAAGCTGCTTGAAAGGGGTAAAGAACTATATGCACTCAAG 840
QY 2424 CGCCAAGTTCTACCGTCAAAAACCTGCCGTATACATCACCAAGAAAGACTTCAAGTTCAAC 2483
D 841 AAGAGGGTTGTAGCGGTGTCACACGAGGAGCTCACAAAACCTGCACATGAAGCTGAAT 900
QY 2484 GACAAAACGGCAAAATACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGCAAA 2543
D 901 GATGCGCTTCCAGAGATCACCGTGGATCCAGAGACATACGTTGTACAGCAAAATGGCGAG 960
QY 2544 CT-CTGCACTCTAAACCCACCTCGCAAGTGCCTCTAGCCGAGCTACACTTTCTTCT 2601
D 961 GTCTTAGCATGTGCCAGCGCATTCAGTCCCTCTCTGCGCGGAACTATTTCTCT 1019

```

```

RESULT 4
LOCUS GR_EB04F01.f GR_Eb Gossypium raimondii mRNA linear EST 16-JUN-2004
DEFINITION mRNA sequence.
VERSION COI22599
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii

```

```

REFERENCE
AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rtwing@genome.arizona.edu
Plate: 04 row: F column: 01.
Location/Qualifiers
1. .872
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb04F01"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site_1: NotI; Site_2: EcoRV; Library made by invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -.80."

```

FEATURES

source

```

Query Match 10.7%; Score 309; DB 7; Length 872;
Best Local Similarity 59.74; Pred. No. 1.3e-65;
Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

```

ORIGIN

```

Query Match 10.7%; Score 309; DB 7; Length 872;
Best Local Similarity 59.74; Pred. No. 1.3e-65;
Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

```

```

QY 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCACCTTCTTCTCCA 1319
D 2 GGAGAAGGATTTGATCGTAACCTGAGGGGCTATTGACTGTGTCATTTTTCATATGTCCT 61
QY 1320 CAACAAATTCCTTACCGCTCTAGCCAATGGGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
D 62 CAGTTGGTTCATGGAAGCTATTTCAAGTGGCATCACAACTATTGAGAGGTGGACCGGA 121
QY 1380 CTTGTAGATGGCA CGAATCGACTACTATCACTCCGGGCAAAATGGAACTTTCACCCGATG 1439
D 122 CCAGCTGAAGGAACACGTGCAACTACTTTGACGCCAGCTCCGTCGCAAAATGAAATGATG 181
QY 1440 TTGCGCGCAGCAGAAAGATTTCTATGAATGTGGCTTTTGGGCAAGGCAATAGCTCT 1499
D 182 CTGCACTCCACTGATGACTTCCCTCTAAATTTTGGCTTCACAGGGAAGGAAATGGTTCT 241
QY 1500 AGCAAAAACAACTTGTAGAACAAAGTAGAAGCGGGCGGATTTGTTTTAAATTCGATGAA 1559
D 242 AAACCTGAAGAAATACATGAAATTAATCAAAAGCCGGAGCAATGGACTGAAACTGCATGAG 301
QY 1560 GACTGGGGCAACACCAAGTGGATCGATCACTGCTTGGAGCGTGGCAGATGAATACGAT 1619
D 302 GATTGGGGAACACCTGCTGCAATAGACAGTTGTTTGGCTGTTCGAGAACTATATGAT 361
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGACAGTTATGTAGATGACACCTA 1679
D 362 ATCCAAGTTTAATATTCATCTACTGACACCTTGAACGAATCTGGAATTTGTGGAACAAAT 421
QY 1680 AATGCAATGAACGGCGCGCATCTCATGCTTACCACATTCAGGAGCGGGTGGAGGACAC 1739
D 422 AATGCAATTTAAAGGAAGAACTATTCACCTTATCACAGTGAAGTGCCTGTTGGTGTGAT 481
QY 1740 TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCCCTCTCCACCAACCC 1799
D 482 GCTCCAGATATCATCAAGTATGTGGCGTTAAAAAAGTCTCCCTTCGTCAACAAACCCG 541
QY 1800 ACTATTCCCTATACCATTAATACGTTGCGAGAACACTTAGACATCTCATGACATGCCAC 1859
D 542 ACTCGCCCTTATATCTTCCAATCTATATAGATGAACATCTTGACATGCTGATGTTGCCAT 601
QY 1860 CACTAGACAAACGCATCCGCGAGGATTTCAATTTTCTCAAAAGCCGCTATCCGCCCGC 1919
D 602 CACTCAGCAAGGATATTCCTGGAAGATGTAGCATTTGCAGATCAAGGATTTAGGCGCGAA 661
QY 1920 TCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGCATGACAAAGTCTCGAT 1979
D 662 ACCATTGCTCCGAGACATATTCATGACACAGGGGCAATCAGCATTTATCTTCAGAT 721
QY 1980 TCGAAGCAATGGGGCGTGCAGCGAAGTATTCCTCGAACTTGGCAGACTGCGGATAG 2039
D 722 TCACAGCGGATGGGTGCGATTTGGCGAGTGATATGCAAGAACTTGGCAAAACCGCCACAAG 781
QY 2040 AATAAAAAGAAATTTGGTAAAGCTTCTTGAAGATGGAAAGATACGATAATTTCCGCATT 2099
D 782 ATGAATATCAAAAGGGGACCACTTGGCTCTGCACTCAGCAACAGCAATCTAAGAAATC 841
QY 2100 AAGCGCTACATCTCCAAATACACTATCAA 2128
D 842 AAACGTTATATAGCAAAATACACCATAAA 870

```

RESULT 5

DR633422

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST

Gibberella moniliformis

Gibberella moniliformis

Eukaryota; Fungi; Ascomycota; Hypocreales; Nectriaceae; Gibberella.

DR633422 925 bp mRNA linear EST 11-JUL-2005
 EST1024047 Fw Gibberella moniliformis cDNA clone FVMA534, mRNA
 sequence.

ACCESSION DR633422
 VERSION DR633422.1
 KEYWORDS GI:70708256
 SOURCE Gibberella moniliformis

ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Hypocreales; Nectriaceae; Gibberella.

1 (bases 1 to 925)
 Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L, Lee,Y.,
 Uterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
 Kendra,D.F., Town,C.D. and Whitelaw,C.A.
TITLE
 Analysis of 87,000 expressed sequence tags reveals alternatively
 spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL
 Unpublished (2005)
COMMENT
 Contact: Brown, D.W.
 USDA/ARS/NCAUR
 1815 N. University St, Peoria, IL 61604, USA
 Tel: 309 681 6230
 Fax: 309 681 6889
 Email: brown@ncaur.usda.gov
 TIGR sequence name: FVMA534TH
 Seq primer: AAT TAA CCC TCA CTA AAG GG.
FEATURES
 source
 1..925
 /organism="Gibberella moniliformis"
 /mol_type="mRNA"
 /strain="m3125"
 /db_xref="taxon:117187"
 /clone="FVMA534"
 /tissue_type="mycelia"
 /clone_lib="FvM"
 /notes="vector: pBluescript II SK(+)"
 1815 N. University St, Peoria, IL 61604, USA
 Site 2: XhoI; anamorph: Fusarium verticillioides. Library
 FvM was prepared from pooled RNA obtained from a 48-hour
 and a 72 hour, liquid GVAM culture from strain M-3125.
 Cultures were vacuum filtered and the mycelial mats were
 frozen in liquid nitrogen, ground to a powder, and then
 added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
 approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
 directionally ligated into the pBluescript II SK(+) XR
 vector (cDNA Synthesis Kit; Stratagene)."
ORIGIN
 Query Match 10.6%; Score 306.2; DB 8; Length 925;
 Best Local Similarity 59.9%; Pred. No. 6.7e-65;
 Matches 548; Conservative 0; Mismatches 363; Indels 4; Gaps 2;
 QY 1312 TTTCTCCACAAATCCCTACCGCTCTAGCCAATGGCTTACAACCATGTTTGGAGCG 1371
 Db 10 TCTGCCCTAGCAAGTACCCGAAGCTTTGCATCTGTGTAAACACCAATGCTTGGCGGTG 69
 QY 1372 GCACAGGTCCTGTAGATGGCAGAAATCGCACTACTATCACTCCGGGCAATGGAATTCG 1431
 Db 70 GTACCGGCCCAAGTGTGTGGAACGAAACGCAACTACTTGTACGCTGTGTCTCATTCATGC 129
 QY 1432 ACCGCATGTTCGGCAGCAGAGAGATTCTATGAATGTGGGCTTTTGGGCAAGGCA 1491
 Db 130 GTCAAATGTTTGCAGGCGTGCAGTCAGCTTCTCTATCAATATTTGTTATTACTGGCAAGGTA 189
 QY 1492 ATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAAGCGGCGGCGATTGGTTTAAAT 1551
 Db 190 ATGATAGTTCTCTGAGGGTCTCGGCATCAGGTCAATGCTGTGTGTGGCTCAAGC 249
 QY 1552 TGCATGAAGACTGGGCAACAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATG 1611
 Db 250 TTCATGAGGACTGGGGTTGCACCTCCTGCTGCTATTGACGCTTGTCTCAGTGTCTGTGATG 309
 QY 1612 ATACCATGTGCAAGTTTGTATCCACACCATCAGTCATATGAGGCAAGGTATGTAGATG 1671
 Db 310 AATTTCGATATTCAATGTCTTATTTCACACTGACACGCTTAACGAGTCTGGCTTTTTCGAAT 369
 QY 1672 ACACCCCTAAATGAACGGGCGGCCCATCATGCTTACCACATTTAGGGAGCGGGTG 1731
 Db 370 CTACGATCGCTGCTTTCNAGAACCGACAAATTCATCTATTCACACAGAGGGTGCAGGAG 429
 QY 1732 GAGGACACTCACCTGATGTTATCACCATGCGACGGCGAGTCAATATTCTACCTCTCTCCA 1791
 Db 430 GTGGCCATGTCGGGATCATCTCCGTGTAGAGATCAAAATGTTCTCCCATCATCGA 489

Qy	1792	CGACCCCGCTATTCCTCTATACCATTAATACGGTTGCGAGAACACTTAGACATGCTCATGA	1851
Db	490	CCAACCTTCAAGACCATTTACACGCAATACTCTCGATGAGCATCTCGATATGCTTTATGG	549
Qy	1852	CATGCCACCACTTAGACAAACGCATCCGCGAGGATTTACAATTTCTCAAAGCCGTATCC	1911
Db	550	TCGTGCATCATCTGTCCAGAATATCCGAGAGATGTAGCCTTCGCGGAGCGGTATTC	609
Qy	1912	GCCCCGGCTCTATCGCGGCTGAAGATGTGCTCATGATATGGGTGTGATCGCATGACAA	1971
Db	610	GTGCTGAACCAATTCGTCTGAGGATGTATACACGACAAAGCGCTATCAGCATGATGA	669
Qy	1972	GCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGAAGTGATTCCTCGAACTTGCGCACTG	2031
Db	670	GCTCTCACTCGCAGGCTATGGGCGGTTGCGGAGAGGTCGTTTTAAGAACAATGGAATACTG	729
Qy	2032	CGCATAGAATAAAAAAGAAATTTGGTAAGCTTCCTGAAGAT--GCACAAGATAACGATA	2088
Db	730	CGCATAAAAATAAGTGCAGAGGGTTGGTTGCCGAGGATGAGGGCACAGGGGCTGATA	7899
Qy	2089	ATTTCGCGCTTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCG	2148
Db	790	ATCGCGGTGTAAACGCTATGTCAGCAAGTATACIATTAACCCAGCTATTGCTCAGGCT	849
Qy	2149	TGACGAGATATATCGGCTCTGTGGAAGAGGGCAAGATCGCGA-CATTGGTGGTGGAAAT	2207
Db	850	TTGGACATGTCATTGGGAGCATTGAGGTTGGAAAGTTTGTGATCTTGTCTTTTCGGAT	909
Qy	2208	CCTGCCCTTTTGGC 2222	
Db	910	CCTGCGTGGTTGGC 924	
RESULT 6			
LOCUS	CKX676965		
DEFINITION	yydd29h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cdna		
SOURCE	clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE UREASE. ;		
ACCESSION	CKX676965		
VERSION	CKX676965.1		
KEYWORDS	GI:57936920		
SOURCE	EST.		
ORGANISM	Strongylocentrotus purpuratus		
REFERENCE	Strongylocentrotus purpuratus		
AUTHORS	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus. 1 (bases 1 to 748) Coffman,J.A., Robertson,A.J., Clifton,S., Pape,D., Hillier,L., Martin,J.J., Wylie,T., Dante,M., Meyer,R., Theising,B., Bowers,Y., Gibbons,M., Ronko,I., Teagareishvili,R., Ritter,B., Kennedy,S. and Wilson,R.		
TITLE	WashU Sea Urchin EST Project		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Dr. James A. Coffman WashU Sea urchin EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu DNA sequencing by: Washington University Genome Sequencing Center Seq primer: -28RpOT High quality sequence stop: 680. Location/Qualifiers 1. . 748 /organism="Strongylocentrotus purpuratus" /mol_type="mRNA" /db_xref="taxon:7668" /clone="ydd29h04" /lab_host="DH108" /clone_lib="Sea urchin EST Lib1" /note="Vector: pCMVSPORT6.1 (Innitrogen) ; Site 1: Not1;		
FEATURES			
source			

Site_2: Smal; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin *Strongylocentrotus purpuratus*, cloned into the vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN		Query Match	10.5%;	Score 302.2;	DB 8;	Length 748;	
		Best Local Similarity	63.1%;	Pred. No. 6.3e-64;			
		Matches 466;	Conservative 0;	Mismatches 273;	Indels 0;	Gaps 0;	
Qy	1648	TCATGAGGAGCGTATATAGATGACACCCCTAAATCAATGACGCGCGCCCATCATG	1707				
Db	9	TCATGAAATCTCTCTGTGTGGAAGATACCAATAGCTCTTTGAGGCGCGACCATCA	68				
Qy	1708	CCTACACATTTGAGGAGCGGGTGGAGGACACTCAGCTGATGTTATCAACATGGCAGCG	1767				
Db	69	CCTATCACTCGAGGGTGCAGGAGGAGGCCACGCCCTGACATCATGAGGTAGTGGCG	128				
Qy	1768	AGCTCAATATTTACCTCTCCACACCCCGCACTATTCCTATACGATTAATACGGTTG	1827				
Db	129	TGCCCAATGTGTTACCTCTGTACCAACCCACCGCCCTTTCACAGTGAACACCATCG	188				
Qy	1828	CAGAACATTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCCGGAGATT	1887				
Db	189	ATGAACACCTGGATATGCTCATGCTGTGCCACCATCTTGACAAAGATCTAAAGGAAG	248				
Qy	1888	TACAATTTTCTCAAAGCGGTATCGGCCCGGCTTATCGCGGCTGAAGATGTCTCCATG	1947				
Db	249	TTGCCCTTTCAGAGTCTCGCATCAGAGCGGAACCATTTGCTGTGAAGATATCTTTCATG	308				
Qy	1948	ATATGGGTGTGATCGGATGACAAAGCTCGGATTCGGAAGCAATGGGGCGTGCAGGGGAAG	2007				
Db	309	ATCTGGGAGCAATCAGTATTTGTGGCTCTCGATTCGCAGGCGCATGGGGCGTGTGGTGGAG	368				
Qy	2008	TGATTTCTCGAATCTGCAGCTGCGATAGATAGATAAAGAAATTTGGTAAAGCTTCTG	2067				
Db	369	TGATCAACCGGTATCATGGCAGCGCAGATGAAGATGAAGATTTTCAGAGGAGAGCTGTGAG	428				
Qy	2068	AGATGGCAAGATGAAGATAATTTTCGCATTAAGCGCTATCATCTCCAAATACACTATCA	2127				
Db	429	AGAAACGGGTGACACGACAACTAAGGGTCAAGGTTACATTTGCCAGTACAGGTCA	488				
Qy	2128	ACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGATCG	2187				
Db	489	ACCGGCGCATTTGCCACCGCATGGGCCATCTCATTTGGCTCTGTACAGTTGGTAAATGG	548				
Qy	2188	CCGACTTGGTGGTGAATCTCGCTCTTTTGGCGTAAACCCCAAAATCGTGATCAAG	2247				
Db	549	CTGACCTGGTTCTCTGGAATCCGGCATTTCTTGGAGCTAAACCTGACCTCATCATCAAG	608				
Qy	2248	GCGTATGGTGGTCTCTCTGAAATGGCGGATTTCAAGCGGTCTGTGCCCACTCCCAAC	2307				
Db	609	GAGGTTACATTTGTTGGGCAAAATGGGCATGGCCAAATGATCATCCGACCCCTGAC	668				
Qy	2308	CGGTTTATACCGGAAATGTTGGGGATCACGGCAAGGGGAAATTTGACACCAAGCATCA	2367				
Db	669	CTGTCAAGACAGGAAATGTTGGTCTTACGGGAAGTCCATCGGTGAGAACTCTGTCA	728				
Qy	2368	CTTTTGTTCAAAGTCGC	2386				
Db	729	TCTTTGTGTCAAAGCAGC	747				

RESULT 7
DR637386
LOCUS
DEFINITION EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA
889 bp mRNA linear EST 11-JUL-2005
ACCESSION DR637386
VERSION DR637386.1
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis

Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 889)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D., and Whitelaw, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR

1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brownw@ncaur.usda.gov
TIGR sequence name: FVMAV95TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES
source
1..889
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAV95"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="vector: pBluescript II SK(+) XR; Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour culture and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad, CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBluescript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

		Query Match	10.4%;	Score 299.2;	DB 8;	Length 889;	
		Best Local Similarity	59.5%;	Pred. No. 3.7e-63;			
		Matches 524;	Conservative 0;	Mismatches 353;	Indels 3;	Gaps 1;	
Qy	1312	TTTCTCCACAACTATCCCTACCGCTCTAGCCAAATGGGTTACACCACTGTTTGAGCGG	1371				
Db	10	TCTGCCCTCAGCAAGTACCCGAACTCTTGATCTGGTGTAAACCACTGCTTGGCGGTG	69				
Qy	1372	GCACAGTCTCTGTATAGGACGAATCGGACTACTATCACTCCGGGCAATGGAACTTGC	1431				
Db	70	GTACCGGCCCAAGTGTGGAAACGAACTACTTGTACGCTGGTGTCTCAATACATGC	129				
Qy	1432	ACCGCATGTTGCGCGCAGCAAGATTTCTATGAATGTGGGCTTTTGGCAAGGCA	1491				
Db	130	GTCAAAATGTTGACGCGTGCATCAGCTTCCTNCAATATTTGTTACTGGCAAGGTA	189				
Qy	1492	ATAGCTCTAGCAAAAAACAACTTGTAGAACAGTAGAAGCGGCGCGATTTGGTTTAAAT	1551				
Db	190	ATGATAGTTCTCTGAGGGTCTGCGCATCAGGTCAATGCTGTGCTTGTGCGCTCAAGC	249				
Qy	1552	TGATGAAAGATGGGGCACAACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATG	1611				
Db	250	TTCAATGAGGACTGGGGTTGCACCTCTGCTGTATTTGACGCTTGTCTAGTGTCTGTGATG	309				
Qy	1612	AATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTATGTAGATG	1671				
Db	310	AATTCGATATTCATGTTCTTATTCACATGACACGCTTAAAGATCTGGCTTTGTGCAAT	369				
Qy	1672	ACACCTCAAAATGCAATGAACGGGCGCGCATCTCCATGCCCTACCAATTTGAGGGAGCGGTG	1731				
Db	370	CTACGATCGCTGCTTTCAGAAACCGCAATTCATCTTATCACACAGAGGGTGCAGGAG	429				
Qy	1732	GAGGACACTCACTGATGTTATTCACCATGGCAGCGAGCTCAATATTCTACCTCTCTCCA	1791				

430	GTGGCCANGTCCGGATATCATCTCCGTGTGAGACATCAAAATGTTCTGCCAATCATCGA	489
1792	CCACCCCACTATTCCTATACCAATTAAATACGGTTGCAGAACACTTAGACATGCTCATGA	1851
490	CCAACTTCAAGACCAATTCACAGCNAATCTCTCGATGAGCATCTCGATATGCTTATGG	549
1852	CATGCCACCACTAGACAAACGGCATCCGCGAGGATTTTCACAATTTCTTCAAAGCGGTATCC	1911
550	TCTTGCCATCACTTGTCCAGAAATATCCAGAGGATGTAGCCTTCGCGGAGGCGGTATTC	609
1912	GCCCCGCTCTATFCGCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCATGACAA	1971
610	GTGCTGAACCAATTGCTGTCTGAGGATGTATTACACGACAAAGGCGTATCAGCATGATGA	669
1972	GCTCGAATTGCGAAGCAATGGGCGTGCAGCGGAAGTGATTCCTCGAACTTGCACAGCTG	2031
670	GCTCTGACTCGCAGGCTATGGGCGGTGCGGAGAGGTGCTTTTAAAGAACATGGAATACTG	729
2032	CGGATGAAGATATAAAAGAAATTTGGTAAGCTTCTGAAAT---GGCAAGATAACGATA	2088
730	CGCATAAAAATAAGGTGCAGAGGGGTGGTTGCCGGAGGATGAGGGCACAGGGGCTGATA	789
2089	ATTTCCGATTTAAGCGGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCG	2148
790	ATCGCGGTGTAAAACGCTATGTGAGCAAGATATACTATTAAACCCAGCTATTGCTCAGGGCT	849
2149	TGAGCAGATATATCGGCTCTGTGGAAGAGGGCAAGATCGC	2188
850	TTGGACATGTCATTGGAAGCATTTGAGGTTGGAAAGTTTGC	889

RESULT 8
 LOCUS DR633939
 DEFINITION EST1024564 FvM Gibberella moniliformis cDNA clone FVMA846, mRNA sequence.
 ACCESSION DR633939
 VERSION DR633939.1 GI:70708773
 KEYWORDS EST.
 SOURCE Gibberella moniliformis
 ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 1 (bases 1 to 796)
 AUTHORS Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D. and Whiselaw, C.A.
 TITLE Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster Unpublished (2005)
 JOURNAL Contact: Brown, D.W.
 COMMENT USDA/ARS/NCAUR
 USDA
 1815 N. University St, Peoria, IL 61604, USA
 Tel: 309 681 6230
 Fax: 309 681 6689
 Email: brown@ncaur.usda.gov
 TIGR sequence name: FVMA846TH
 Seq primer: AAT TAA CCC TCA CTA AAG GG.
 Location/Qualifiers
 1..796
 /organism="Gibberella moniliformis"
 /mol_type="mRNA"
 /strain="m3125"
 /db_xref="taxon:117187"
 /clone="FVMA846"
 /tissue_type="mycelia"
 /clone_lib="FvM"
 /note="vector: pBlueScript II SK(+) XR; Site_1: EcoRI; Site_2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour culture and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were
 FEATURES
 source

RESULT 9
DR635599
LOCUS
DEFINITI

ACCESSION DR635599
VERSION DR635599.1 GI:70710433
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
AUTHORS 1 (bases 1 to 849)
Brown, D.W., Cheung, F., Proctor, R.H., Butcho, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.
TITLE Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL Unpublished (2005)
COMMENT Contact: Brown, D.W.
USDA/ARS/NCAR
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FVMAJ96TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
FEATURES
source
1..849
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAJ96"
/tissue_type="mycelia"
/clone_lib="FVM"
/note="Vector: pBluescript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FVM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBluescript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."
ORIGIN
Query Match 10.1%; Score 292; DB 8; Length 849;
Best Local Similarity 60.2%; Pred. No. 2.3e-61;
Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;
QY 1317 CCACAAATTCCTACCGCTCTAGCCCAATGGCGTTTACACCACTATTTGGAGGCGGCACA 1376
DB 3 CCTCAGCAAGTACCCGAAGCTCTTGCACTCTGCTTAACCACTGCTTGGCGGTGTACC 62
QY 1377 GGTCTGTAGATGGCAGATGGCACTACTATCACTCGGGCAATGGAACTTGCACGC 1436
DB 63 GGCCCCAAGTCTGGAACGACGCAACTACTTGTACGCTGTGTCTCAATCATGCGTCAA 122
QY 1437 ATGTTGGCGCAGCAGCAAGAGTATTTCTATGAATCTGGCTTTTGGCAAGGCAATAGC 1496
DB 123 ATGTTGGCGCGTGGATCAGCTTCCTTCAATATGTTATCTGCAAGGTATGAT 182
QY 1497 TCTAGCAAAAAACAACTTTGTAGAACAGAGTGAAGCGGCGGATTTGTTTTAAATTCAT 1556
DB 183 AGTTCTCTCAGGGTCTGCGGATCAGCTCAATGCTGTGCTTGTGGCTCAAGCTTCAT 242
QY 1557 GAAGACTGGGCAACACCAAGTGGCATGCATCACTGCTTGAGCTGGCGAGATATAC 1616
DB 243 GAGGACTGGGTTTGCACTCTCTGCTATTGACGCTTGTCTCACTGTCTGTGATGAATTC 302
QY 1617 GATGTGAAGTTTGTATCCACCGCATACATCAATGAGGCGAGTATGTAGATGACACC 1676
DB 303 GATATTCAATGTCTTATTCACACTGACACGCTTAACGAGTCTGCTTTGTGCAATCTACG 362
QY 1677 CTTAATGCAATGAACGGCGCGCCATCCATGCCCTACCACTTTGAGGAGCGGTGGAGGA 1736

DB 363 ATCGCTGCTTCAAGAACCGCAATTCATCTTATCACACAGAGGTGCAGAGGTGGC 422
QY 1737 CACTCACCTGATGTTATCACCATGGCAGGAGCTCAATATTCTTACCTCTCCACCAACC 1796
DB 423 CATGCTCCGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAAC 482
QY 1797 CCACACTATCCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGC 1856
DB 483 CTTACAAGACCATTTCAACGCAATACTCTCGATGAGCATCTCGATATGCTTATGGTCTGC 542
QY 1857 CACCACCTAGACAAACGCACTCCGCGAGGATTTCAAAATTTCTCAAAGCCGCTATCCGCCCC 1916
DB 543 CATCACTTGTCCAAGATATATCCAGAGSATGTACGCTTCCGCGAGAGCCGTATTCGTCT 602
QY 1917 GGCTCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAAAGCTCG 1976
DB 603 GAAACCATTCGCTGCTGAGGATGATTACACGACAAAGGCGCTATCAGCATGATGAGCTCT 662
QY 1977 GATTCGCAAGCAATGGGCGTGCAGGCGAAGTATTCTCGAACTTGCAGACTGCGGAT 2036
DB 663 GACTCGCAGGCTATGGGCCGTTGCGGAGAGTCTGTTTAAAGAACATGGAATCTGCGCAT 722
QY 2037 AAGAATAAAAAAGAAATTTGTTAAGCTTCTCTGAAGAT---GGCAAGATAACGATAAATTC 2093
DB 723 AAAAATAAGTGCAGAGGGGTTGTTCCGCGAGATGAGGCGACAGGGGCTGATTAATCGG 782
QY 2094 CGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGGCGT 2149
DB 783 CGGTAAACGCTATGTGACCAAGTATCTATTAAACCCAGCTATTGCTCAGGGCTT 838
RESULT 10
CV195867
LOCUS 766 bp mRNA linear EST 15-SEP-2004
DEFINITION CGF1003441.1 All Seed coat from mid-season walnut embryos collected
Aug 1 Juglans regia cDNA clone WSC0002_IVF_A11 5', mRNA sequence.
VERSION CV195867
KEYWORDS CV195867.1 GI:52124704
SOURCE EST.
ORGANISM Juglans regia (English walnut)
Juglans regia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fagales; Juglandaceae; Juglans.
REFERENCE Muir, R., Baek, J., Leslie, A., Cook, D. and Dandekar, A.
1 (bases 1 to 766)
AUTHORS Analysis of genes expressed in walnut seed coat tissue
TITLE Unpublished (2004)
JOURNAL Contact: Abhaya Dandekar, PhD
COMMENT CAES Genome Facility
UC Davis, Department of Pomology
One Shields Ave. Davis, CA 95616, USA
Tel: 530 752 7784
Fax: 530 752 8502
Email: andandekar@ucdavis.edu
Seq primer: WSCF-TCCGAGATCTGGACGAGC.
FEATURES
source
1..766
/organism="Juglans regia"
/mol_type="mRNA"
/cultivar="Tulare"
/db_xref="taxon:51240"
/clone="WSC0002_IVF_A11"
/sex="Hermaphrodite"
/dev_stage="Mid season fruit collected Aug 1"
/lab_host="XLI0-Gold"
/clone_lib="Seed coat from mid-season walnut embryos
collected Aug 1"
/note="Organ: Seed coat; Vector: pTriplex2; Site 1: SfIIA;
Site 2: SfIIB; Walnut nut samples were harvested from
Tulare trees growing in the 'Stuke Block' in the Wolfskill
experimental orchard located in Winters, California (USA)."

Samples were harvested on August 1, 2001 between 8 and 10 am. Samples were then dissected the same day. Seed coat (pellicle) tissue was separated from embryos and frozen immediately in liquid nitrogen and stored at -80C. A gram of sample was removed and ground to a fine powder in liquid nitrogen. Total RNA was extracted using the hot borate procedure. Poly A+ RNA was obtained using the Poly(A) Purist kit (Ambion). The cDNA Library was constructed using the SMART cDNA library Kit (Clontech). Primary library was en masse evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (XL10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/tray), picked using a Qbot and archived in 384 well dishes."

ORIGIN

Query Match	9.7%	Score 279.6	DB 7	Length 766
Best Local Similarity	60.3%	Pred. No. 2.8e-58		
Matches 462	Conservative 0	Mismatches 304	Indels 0	Gaps 0
QY	1368	GGCGGCACAGGTCCTGTAGATGCGACGAATGCGACTACTACTACTCTCGGSCAAATGGAAC	1427	
DB	1	GGGGGACGGGACCTGCTGATGGAACGGTGGCAACACTTGTACTCAGCAATATCAAA	60	
QY	1428	TTGCACCGCATGTTGGCGCGACGAGAGATATTTCTATGAATGGGCTTTTTGGGCAAA	1487	
DB	61	ATGAAGTTAATGCTGCAATCAACTGATGAGATGCCTCTAAATTTTGTTTTCAGGGAAA	120	
QY	1488	GGCAATAGCTCTAGCAAAAACAACTTTGTAGAACCAAGTAGAAGCGGGCGGATTTGGTTTT	1547	
DB	121	GGGAACAGTGC AAAA CCTGTATGAACACTACATGAATATATTAGAGCCGGGCAATGGGACTG	180	
QY	1548	AAATTGCATGAAGACTGGGGCAACAAACCAAGTGCAGTCGATCCTCTTTGAGCGTGCCA	1607	
DB	181	AAGCTGCATGAGACTTGGGGAACCTACTCTCTGCTGCAATAGACAATGCTTTGAATGTTGCA	240	
QY	1608	GATGAATACGATGTGCAAGTTTGTATCCACACCGATACTGTCATGATGAGGCGAGTTATGTA	1667	
DB	241	GAACAAATATGACATCCAGGTTAATATATCCACACGACACTTGAATGAATCGGGATTTGTT	300	
QY	1668	GATGACACCCATAATGCAATGAACGGGCGCGCCATCCATCGCTTACCACATTTGAGGGAGCG	1727	
DB	301	GAACATTTATGCTGCAATTTAAAGGAAGAACTATTTCATACCTACCACTGAAGGTGCT	360	
QY	1728	GGTGGAGGACATTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTACCCCTCC	1787	
DB	361	GGTGGTGGTCATGCTCCAGATAATCAATTAAGTATGGTGTAAACATGCTCTGCCATCA	420	
QY	1788	TCACACACCCCACTATTTCCCTATACCATTAATACGGTTGCGAGAACACTTAGACATGCTC	1847	
DB	421	TCACGAACCCCAACAGCGCCCTTTTACTTTCAACACTATAGATGAGCATCTTGACATGCTG	480	
QY	1848	ATGACATGCCACCACTAGACAAACGATCCGGAGGATTTACAAATTTTCTCAAAAGCGT	1907	
DB	481	ATGGTGTCCCATCACCTCGATAAAGATATTTCCAGAAGACGTTGCTTTTGCTGAATCAAGG	540	
QY	1908	ATCCGCCCGCGCTCTATCGCGCTGAAGATGTGCTCCATGATGGGTGTGATCGCGATG	1967	
DB	541	ATAAGGCTGAACAATTTGCTGCAGAGGATATTTTGCATGATATGGGGCAATTAGCATC	600	
QY	1968	ACAAGCTCGATTTCGCAAGCAATGGGCGGTGACGGCGAAGTGATTCTCGAATTGGCAG	2027	
DB	601	ATTGCTTCCGATGCAACAGCTATGGGTTCGCAATTTGGAGAGGTGATTAGCAAACTTTGGCAA	660	
QY	2028	ACTGCGGATAAGAAATATAAAGAAATTTGGTAAAGCTTCTCGAAGATGCCAAAGATAACGAT	2087	
DB	661	ACTGCCCAAGATGAAGTTTCAAGAGGGTCAATTGACCCCTAGTGACACGACATGAC	720	
QY	2088	AAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCG	2133	
DB	721	AACTCTTCGTATCAAGGCTTACATTTGCTTAATAACAAATAAATCTCTG	766	

RESULT 11	CF826586/c	LOCUS	DEFINITION
-----------	------------	-------	------------

CF826586 986 bp mRNA linear EST 01-APR-2004
EST703968 *Coccidioides posadasii* saprobic phase cDNA library, 2 to
4 kb *Coccidioides posadasii* cDNA clone C1DB361 3' end, mRNA
sequence.

sequence.
ACCESSION CF826586
VERSION CF826586.1 GI:45932643
KEYWORDS EST.
SOURCE Coccidioides posadasii

REFERENCE 1 (bases 1 to 986)

AUTHORS Gardner, M. V. and Cole, G. T.
TITLE Analysis of gene expression in *Coccidioides posadasii* mycelia and spherules via expressed sequence tags
JOURNAL Unpublished (2003)

JOURNAL	COMMENT
Unpublished (2003)	Other ESTs: EST703969
spineurines via expression	

Other_ESIB: ESI1703369
Contact: Gardner MJ

Contact: Guilaner MD
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 3519

Fax: 301 838 0208

Email: gardner@tigr.org.

Location/Qualifiers

1. .986

/organism="Cocci

/mol_type="mRNA"

```

/strain="C735"

```

/db_xref="taxon:

```
/clone="CIDB361"
```

```
/dev_stage="saprobic phase (mycelia)"
```

```
/lab_host="E. coli DH10B, T1 phage resistant"
```

```
/clone_lib="Coccidioides posadasii saprobic phase cDNA  
library" 2 to 4 kb"
```

ORIGIN

		Query Match	9.5%; Score 274.2; DB 6; Length 986;	
		Best Local Similarity	59.1%; Pred. No. 6.6e-57;	
		Matches 512; Conservative	0; Mismatches 343; Indels 12; Gaps 2;	
QY	1721	GGGAGCGGTGGAGGACACTCACCTGATGTATCACCATGCGAGCGGAGCTCAATATTCT	1780	
DB	986	GGGTGCTGGAGGAGCCGCCTCAGATATCATTCGGTGGTGGAAGCCAACGTCCT	927	
QY	1781	ACCTCTCTCCACCACCCCACAATACTCCCTATACCATTAATACGGTTTCGAAACACATTAGA	1840	
DB	926	GCCAGGAGTAGAATCCCACTCGTCGGTATACGGTAATACTTTAGATGAACATCTGGA	867	
QY	1841	CATGCTCATGACATGCCACCTAGACAAAACGATCCGCGAGGATTTACAAATTTCTCA	1900	
DB	866	CATGGTAATGGTCTGCCATCATTTGTCCAAAGATAITTCCTGAAGACGTGGCTTTTGGCGA	807	
QY	1901	AAGCGGTATCCGCCCCGGCTCTATCCGGGCTGAAGATGTGCTCCATGATATGGGTGTGAT	1960	
DB	806	AAGCCGGATCCGATCCGAGACAAATTCGTGCAAGAAGCGTTCTTCATGACACGGGAGCCAT	747	
QY	1961	CGCGATGACAAGCTCCGATTTCGAACCAATGGGGCGTCAGGGCGAAGTGATTCTCTCGAAC	2020	
DB	746	CAGCATGCTATCTCCGACTCTCAAGCTATGGAACGCTGTGGAGAAGTTGTTGTTCCGAC	687	
QY	2021	TTGGCAGACTCCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTTCCTGAAGAT---GGCAA	2077	
DB	686	ATGGAACATGACATAAGATAAATAATGGAACGAGGGCGACTCAAGGAAGATGAAGGAC	627	
QY	2078	AGATAACGATAATTTCCGCATTGAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTT	2137	

Db 626 GGATTCCTGATAATTTTAGGGTTAAACGGGTATATACAGCAAGTACACCATCAACCCCTGCCAT 567

QY 2138 GACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGGGCAAGATCGCCGACTTGGT 2197

Db 566 TGCACAGGGGATGGCCCACTATTTGGAGCGTGGAAAGTTGGCAAGACCGCTGATTGGT 507

QY 2198 GGTGTGGAATCTGCGCTTTTGGCGGTAAACCCAAATCTGTGATCAAAAGGCGGTATGGT 2257

Db 506 TCTGTGGAATTTGCCAACTTTTGGGACTAAACCGAGTATGGTCTTGAAGTCTCGAATGGC 447

QY 2258 GGTCTTCTGTAATGGCGATTTCAACGGCTGTGTCGCCACTCCCCACCGGTATTATTA 2317

Db 446 TGTCTCAGCGCAGATGGGTATCCCAATGGCTCTATCCCAATCGAGCTTATTAT 387

QY 2318 CCGGAAATGTTTGGGCATCACGGCAAGGCAAAATTTGACACAGCATCACTTTTGTTC 2377

Db 386 GAGGCTATGT-----ACGTAGTCTTAACCTTAAGCCTCAATCATGTTCGTATC 336

QY 2378 CAAAGTCGCTATGAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAGTTCCTACC 2437

Db 335 CCAAGCATCCATCAAGCTTGGTATCATCGACAGTTACCACTGAAGAAGCGGATCGAGCC 276

QY 2438 GGTCAAAATCGCGTAAACATCAACCAAGAGCTTCAAGTTCAACGACAAACGGCAA 2497

Db 275 AGTGAAGAATTTGCGGAATATAAGCAAGAGATATGAATTTAATGATATTATGCCAA 216

QY 2498 AATCACCCTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAA 2557

Db 215 AATGAGAGTCGATCCGAGAGCTATGTTGTCGAGGCTGACGGGAGAGTGCACCGCTGA 156

QY 2558 ACCACCTCGCAAGTGCCTTAGCCCCA 2584

Db 155 GCGAGTCAGAGTTCCTTTAAACACA 129

RESULT 12

AZ935182/c

LOCUS BJ_Ba0003020f B. japonicum BAC library Bradyrhizobium japonicum genomic, genomic survey sequence.

DEFINITION

ACCESSION AZ935182.1 GI:13777494

VERSION

KEYWORDS

SOURCE

ORGANISM

Bradyrhizobium japonicum

Bradyrhizobium japonicum

Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

Bradyrhizobiaceae; Bradyrhizobium.

Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,

Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.

A marker-dense, sequence-ready map of the Bradyrhizobium japonicum genome

Genome Res. 11 (8), 1434-1440 (2001)

11483585

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Class: BAC ends

High quality sequence stop: 719.

Location/Qualifiers

1. 784

/organism="Bradyrhizobium japonicum"

/mol_type="genomic DNA"

/strain="USD110"

/db_xref="taxon:375"

/lab_host="E. coli"

/clone_lib="B. japonicum BAC library"

/note="Vector: pindigo536; Site_1: HindIII"

FEATURES

source

1. 784

/organism="Bradyrhizobium japonicum"

/mol_type="genomic DNA"

/strain="USD110"

/db_xref="taxon:375"

/lab_host="E. coli"

/clone_lib="B. japonicum BAC library"

/note="Vector: pindigo536; Site_1: HindIII"

ORIGIN

Query Match 9.5%; Score 273; DB 9; Length 784;

Best Local Similarity 64.0%; Pred. No. 1.2e-56;

Matches 458; Conservative 0; Mismatches 253; Indels 5; Gaps 3;

QY 1468 ATCTGGCTTTTGGCCAAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAAAGTAG 1527

Db 715 ATCTGGCATTTTCGGGCAAGGGCAACGCCCTCGCGCCCGCGCCCGCTGGTCAGATGATCA 656

QY 1528 AAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAACCAACCAAGTGCATCG 1587

Db 655 AGNCGGCGCATGCGCGCTGAAGCTGCAAGAGGATTGGGGCAC-ACGCGGCGCGATCG 597

QY 1588 ATCACTGCTTTAGCGTGGCAGATGAATAAGATGCAAGTTTGTATTCACACCGGATACAG 1647

Db 596 ACAATGCTGTGGTGGCGGACGATTTACGACATCCAGGTCAATGATCCACACCGATAGCG 537

QY 1648 TCAATGAGGAGGTTTGTAGATGACACCTTAATGCAATGAAGCGGGCGGCATCCCATG 1707

Db 536 TGAAGATCCGGGCTTCGTGAGGATACGATCAAGGGGTTCAAGGGCGCGCACCATCCACG 477

QY 1708 CCTACCACTATGAGGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCG 1767

Db 476 CTTTCCACACCGAGGGCGCGCGCGGTCAAGGCGGCTTCAAGGCGCGCGCGCGCGCGCG 417

QY 1768 AGCTCAATATTTCAACCTCTCTCCACACCCCACTATTTCCCTATACCAATTAATACGGTTG 1827

Db 416 TGAAGAACGTGCTGCCGTCTCGACCAACCGAGCGCGCCCTTCAAGGCGCGCGCGCGCGCG 357

QY 1828 CAGAACATTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATT 1887

Db 356 ACAGCATCTCGACATGCTGATGTTGTGCCACCACTCGATCCCTCGCATCGCGGAAGATC 297

QY 1888 TACAATTTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATG 1947

Db 296 TTGGTTTCGCGAAGCCGTATCCGCAAGAGACCATCGCGCGCGAGGACATCCTGCGACG 237

QY 1948 ATATGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGG-GCGTGCAGGCGAA 2006

Db 236 ATCTCGCGCGCTCTCGATGATGCTCGTCTCGNACTCCAGGCCATGGGCGCGCTCGCGAA 177

QY 2007 GTGATTCCTCGAATTGGCAGATCGCGGATGAAGATTAAGAAAGATTTGGTAAAGTTCC- 2065

Db 176 GTCATCATCGGACCTTCGCGAGACCGCGGACAGATGAAGATGAAGACGCGCGGATCGCTGCG 117

QY 2066 --TGACATCGCAAGATGAAGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT 2123

Db 116 CAGGCAAGGCGAAGGACAGCAATTTCCGCGTCAAGCGCTACATCGCCAAATACAG 57

QY 2124 ATCAACCCCGCTTTGACCCACCGCGTGAAGGAGTATATCGGCTCTGTGGAAGAGGG 2179

Db 56 ATCAACNCGGATCGCGCAGCGGCTGTGGAAGCTGATCGGTTCGGTGGAGAAGG 1

RESULT 13

CO027797/c

LOCUS

DEFINITION

EST806181 Coccidioides posadasii spherule cDNA library, 0.5 to 5.3 Kb Coccidioides posadasii cDNA clone CIPAM57 3' end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Coccidioides posadasii

Coccidioides posadasii

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

Onygenales; mitosporic Onygenales; Coccidioides.

REFERENCE

1 (bases 1 to 990)

AUTHORS

Gardner,M.J. and Cole,G.T.

TITLE

Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags

JOURNAL

COMMENT

Unpublished (2003)

Other ESTs: EST806182

Contact: Gardner MJ

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igr.org

FEATURES

Location/Qualifiers

source

1..990
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIPAM57"
/dev_stage="spherules"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii spherule cDNA library,
0.5 to 5.3 kb"
/notes="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV;
Coccidioides posadasii spherule cDNA library, 0.5 to 5.3
kb"

ORIGIN

Query Match 9.1%; Score 262.6; DB 7; Length 990;
Best Local Similarity 58.8%; Pred. No. 5.2e-54;
Matches 525; Conservative 0; Mismatches 339; Indels 29; Gaps 3;

1696 GCGCCATCCATGCTACCACTTGGAGGAGCGGTGGAGGACACTCACCTG-ATGTTATC 1754
990 GAGTGATTACATGCTACCACTGAGGTGCTGGAGGAGCCAGCTCCAGATATCATY 931
1755 ACCATGGAGCGAGTCAATATTCACCTCTCCACCAACCCCACTATTCCTATACC 1814
930 TCCGTGCTGAGAAGCAACGCTCTGCCAGCAGTACGAATCCCACTCGTCGGTATACG 871
1815 ATTAACTCGTTTCAGAACACTTAGACATGCTACATGACATGCCACCACTAGACAAACGC 1874
870 GTAATACTTTAGATGAACATCTGGACATGGTAATGCTCTGCCATCATTTGTCCAAAGAT 811
1875 ATCCGCGAGGATTTACAAATTTCTCAAAGCGGTATCGGCCCGGCTCTATCGCGGCTGAA 1934
810 ATTCTGAGACGTGGCTTTTGGGAAGCGCGATCCGATCCGAGACATTTGCTGCAGAA 751
1935 GATGCTCTCATGATATGGGTGATGTCGGATGACAAAGCTCGGATTCGCAAGCAATGGGG 1994
750 GACGTTCTTCATGACAGCGGAGCATCAGCATGCTATCTCTCGACTCTCAAGCTATGGGA 691
1995 CGTGCAGGCGAAGTATTCCTCGAACTTGGCAGACTGGGATGAAGATTAAGAAAGATTT 2054
690 CGCTGTGGAGAAGTTGTTGTCGACATGGAACACTGACATAGAAATTAAGATGGAACGA 631
2055 GGTAAAGCTTCTCAAGAT--GGCAAGATTAACGATAATTTCCGCAATTAAGCGCTACATC 2111
630 GGGCGACTCAAGGAAGATGAGGAGCGGATTTGATAAATTTTAGGGTTAAACGGTATATC 571
2112 TCCAAATACATATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTG 2171
570 AGCAAGTACACCATCAACCTGCAATGACAGGGGATGGCCACACTATTTGGAGCGTG 511
2172 GAAGAGGCAAGATCGCGACTTGGTGTGTGGAATCTCTGCTTTTGGCGTAAACCC 2231
510 GAAGTTGGCAAGACCGCTGATTTGGTTCTGTGGAATTTGGCAACTTTGGGACTAAACCG 451
2232 ABAATCGTATCAAGGCGGTATGGTGTCTCTGTAATGGCGGATTTCTAACGGCTCT 2291
450 AGTATGGTCTTGAAGTCTGGAATGGCTGTCTCAGCGCAGATGGGTGATCCCAATGGCTCT 391
2292 GTCCCACTCCCAACCGGTTTATTACCGGAAATGTTTGGGCGATCAGCGGAGGCGAA 2351
390 ATCCCAACATCGAGGCTATTATATGAGGCTATGTACGCTCTCA----- 344
2352 TTTGACACCGACATCACTTTGTTTCCAAAGTGGCCCTATGAAATATGGCGTGAAGAAAG 2411
343 -----ATCATGTTCTGATCCCAAGCATCCATCAAGCTTGGTATCATCGACAGT 296

2412 CTGGGCTTAGAGCCCAAGTTCTTACCGGTCAAAATCGCGTAACATCACCAGAAAGAC 2471
295 TACCACCTGAAGAAGCGGATCGAGCCAGTGAAGAAATTTGTCGGAATATAAGCAAGAGAT 236
2472 TTCAAGTTTCAACGACAAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTT 2531
235 ATGAATTTAATGATATTATGCCCATAATGAGATCGATCCGGAGAGTATTTGTTCGAG 176
2532 GTAGATGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTCTAGCCCCA 2584
175 GCTGACGGGAGAGTGCACCGCTGAGCCAGTGTACAGATTTGCCTTTAAACACA 123

RESULT 14
CN907928
LOCUS
DEFINITION
0310109ABLC001286HT (ABLC) Braeburn cell culture three days after
subculture Malus x domestica cDNA clone ABLC001286, mRNA sequence.
CN907928
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Malus x domestica
Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 684)
AUTHORS
Bauming, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,
McArtney, S., Newcomb, R., Ross, G., Snowden, K., Walton, E. and Yauk, Y.
TITLE
JOURNAL
COMMENT
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave, A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.

FEATURES
source
1..684
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABLC001286"
/tissue_type="Cell cultures"
/dev_stage="Exponential phase of growth (three days after
subculture)"
/clone_lib=" (ABLC) Braeburn cell culture three days after
subculture"
/notes="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

ORIGIN

Query Match 9.1%; Score 261.2; DB 7; Length 684;
Best Local Similarity 61.4%; Pred. No. 1.1e-53;
Matches 419; Conservative 0; Mismatches 263; Indels 0; Gaps 0;

1546 TTTAAATTGCATGAAGACTGGGGCACAACACCAAGTGGGATCGATCCTGTTGAGCGTGG 1605
3 TGAAGCTGCATGAGGACTGGGGAAGCACTCTGCTGCAATCGCAATTTGTTTGTCTGTG 62
1606 CAGATGAATACGATGTGCAAGTTTGTATCCACCCGATACAGTCAATGAGCGAGTTATG 1665
63 GAGATCAATATGACATCCAGGTTAACTATCCATCCATAGACACCTTGAATGAATCTGGATTG 122
1666 TAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCATGCTTACCATTTAGGGAG 1725
123 TAGAGCATACAATTGCTGCATTTAAGGAAGAACTATCCATACTTACCACAGTGAAGTG 182
1726 CGGGTGGAGGACACTCACTGATGTTATACACCTGAGCGGAGCTCAATATTCTACCT 1785
183 CAGGTGGGGGCGCATGCTCCAGATATAATCAAGTCTGCGGTGTAAGAAATGTCCTGCCAT 242

1786 CCTCCACCCACCCACTATTCCTTATACCAATTAATACGGTTGCAGAACACTTACATGCG 1845
1787 |||||
243 CATCAACGAATCCCAACGGCCCTTTCACCTCAAAATACATATAGATGAGCATCTTGTATGCG 302
244 |||||
1846 TCATGACATGCCACCACTAGACAAAGCGATCCGCGAGGATTTTACAAATTTCTCAAGCC 1905
1847 |||||
303 TGATGTTTGGCATCACCTTGACAGGACATCCAGNAGATGTAGCTTTTGCTGAATCAA 362
304 |||||
1906 GTATCCGCCCCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTATCGCGA 1965
1907 |||||
363 GAATAAGGGCTGAACAAATTTCTGACAGAGATATTTTGCACGATATGGGGGCCATTAGCA 422
364 |||||
1966 TGACAGCTCGGATTCGCAAGCAATGGGGCTGCAGCGCAAGTGTATTCCTCGAACTTGGC 2025
1967 |||||
423 TTGTATCTTCTGATTCACAGGCTATGGGTGCGAATGGAGAGGTGTAATCAGAACTTGGC 482
424 |||||
2026 AGACTCGGATGAAGATTAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAACG 2085
2027 |||||
483 AAGAGCTGACAGATGAATACAAAGAGGCTCGATAGAACCTAGTGGATCCGACATG 542
484 |||||
2086 ATAATTTCCGATTAAGCGCTACATCTTCAAAATACACTATCAACCCCGCTTTTGACCCACG 2145
2087 |||||
543 ACAATTTCCGATCAACAGATACATCGCAAAGTACACTATAAATCCAGCAATAGTTAATG 602
544 |||||
2146 GCGTAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGTGTGGA 2205
2147 |||||
603 GATTTTCTCATGATGTGGTCTTATTGAGTGGGAGAGCTGGCTGATCTTTGTTCTATGGA 662
604 |||||
2206 ATCCTGCTTTTGGGGTAAA 2227
2207 |||||
663 AGCAATGTTCTTTGGGCGAAA 684
664 |||||

RESULT 15
CF714848/c
LOCUS CF714848 921 bp mRNA linear EST 16-AUG-2004
DEFINITION C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCAF343, mRNA sequence.
ACCESSION CF714848
VERSION CF714848.1 GI:41569007
KEYWORDS EST.
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
ORGANISM Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
REFERENCE 1 (bases 1 to 921)
Loftus, B.
AUTHORS End sequencing of clones from a Full length enriched, normalized
TITLE JEC21 cDNA library
JOURNAL JEC21 cDNA library
COMMENT Unpublished (2003)
Other_ESTs: CCAF343TR
Contact: Brendan Loftus
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: bjoftus@tigr.org
Seq primer: TF.
Location/Qualifiers
FEATURES
source
1. .921
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
/db_xref="taxon:40410"
/clone="CCAF343"
/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"

ORIGIN

Search completed: November 29, 2005, 07:20:55
Job time : 10217 secs

Query Match 9.1%; Score 261; DB 6; Length 921;
Best Local Similarity 55.4%; Pred. No. 1.3e-53;
Matches 504; Conservative 0; Mismatches 405; Indels 0; Gaps 0;
QY 1694 GCGCGCATCCATGCGCTTACCACATTTAGGGAGCGGGTGGAGGACACTCACCTGATGTTAT 1753
DB 921 GAGACATATTCACAGTTACCACTGAGAGAGCTGTGTGGGACATGCTCCAGACATCAT 862
QY 1754 CACCATGGCAGGGAGCTCAATATTTACCTCTCTCCACCAACCCCACTATTCCTATPAC 1813
DB 861 CGTTGTCTGTGAATATGAAAACGCTTTGCCCTAGTTCCACCAACCACTAGGCTTATGC 802
QY 1814 CATTATACGGTTGCGAAGACACTTAGACATGCTCATGACATGCCACCACTAGACAAACG 1873
DB 801 TGTTAAACACTTTTGATGAGCATCTTGACATGCTTATGTTGTGTGCCACCTCGATAAGTC 742
QY 1874 CATCCCGGAGGATTTCAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATTCGGGGCTGA 1933
DB 741 TATCCCGGAGGACANTGCCCTTTGCCGACTCTCGTATCCGTTCTGAAACCGTTGACGCCGA 682
QY 1934 AGATGTCTCCATGATATGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGG 1993
DB 681 AGAGTTTTTACAGGACACAGCGCGATTTCCATGATCTCATCTGACTGTCAAGCTATGGG 622
QY 1994 GCGTGACGGCAAGTGAATCTCGAACTTGGGAGACTGCGGATAGGAATAAAAAAGAAAT 2053
DB 621 TCGTATCGGTGAATCATCATCTGTATACGCGGTACGGCTGCGAAGATGAAGCAATTCGG 562
QY 2054 TGTTAAGCTTCTGAGATGCGAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTC 2113
DB 561 TGGTCTCTTTGAGGGCGATGAACCCAGGAGACAAATAGGGGTGAACCGTATGTTGGC 502
QY 2114 CAATAACTATCAACCCCGCTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGGA 2173
DB 501 CAAGTACACTATTAACCCCGCATTTACCAAGTATGTGCGACCTCATCGGTCAAGTTGC 442
QY 2174 AGAGGGCAAGATCGCGACTTGGTGTGGAATCTTGCCTTTTGGCGGTAAACCCAA 2233
DB 441 CGTGGGCTGTCTGGCAGATCTGGTCTTCTGGAACGGCAGAGTCTTTGGTGTAGACCGGA 382
QY 2234 AATCGTATCAAGGGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAAGCGTCTGT 2293
DB 381 GATGATCTCAAGGGTGTGTCTATGCTGGGCTGCAATGGGCGATGCGAACGCTTCGAT 322
QY 2294 GCCCACTCCCCAACCGGTATTTATACCGGAAATGTTGGGCGATCAAGGCAAGCGAAAT 2353
DB 321 CCCCACTGTGAGCCCGCTCATTTGCGAGCCCATGTGGGGCTCTCAGCCTGAGGCGCTGC 262
QY 2354 TGACACCAAGCATCACTTTTGTTCCTCAAGTGCCTATGAAATGGCGTGAAGAAAGCT 2413
DB 261 ACTCAATTTCAATTTTGGGTGAGCCAGGATCTCTTGACAAGGATCTCGTGAAGAGATT 202
QY 2414 GGGCTTAGAGCGGCAAGTTCTACCGGTCAAAAACCTGCGGTAACTACCAAGAAAGACTT 2473
DB 201 CAACATCAAGAAGAGGGCGGAGGCTGTCAAGAACTGTGCTCAATTTGAAAGAGGATAT 142
QY 2474 CAAGTTCAAGACAAAACGCGCAAAAATCACCGTCGATCCGAAACCTTTCGAGGTCTTTGT 2533
DB 141 GAAGTGGAAATGACAGATGCGGAAGATGACTGTGATCCGGAGACTTACAGCTTACGC 82
QY 2534 AGATGGCAAACTCTGCACCTCTAAACCCACCTCCGAGTGGCTCTTAGCCCGAGCTACAC 2593
DB 81 TGACGGGCTCTTTGCGACGCTCCACCGGACAGACAACTCCCACTGACCAAGAGATCTT 22
QY 2594 TTTCTTTCTA 2602
DB 21 CGTTTACTA 13

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:04:09 ; Search time 522 Seconds
(without alignments)
9817.453 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 rrgagattttccarctt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1075.2	37.3	2735	2	US-08-920-095-1
2	1075.2	37.3	2735	6	PCT-US96-05800-1
3	1062.6	36.9	2619	2	US-08-467-822-19
4	1062.6	36.9	2619	3	US-08-432-697-19
5	1062.6	36.9	2619	3	US-08-466-248-19
6	1058.6	36.7	4824	3	US-09-431-705-1
7	1058.6	36.7	4824	3	US-09-431-705-19
8	737.6	25.6	6131	2	US-07-732-242C-8
9	693.4	24.1	2400	2	US-08-967-513-1
10	693.4	24.1	2400	2	US-08-687-645B-1
C 11	664.4	23.0	1830121	3	US-09-557-884-1
C 12	664.4	23.0	1830121	3	US-09-643-990A-1
C 13	664.4	23.0	1830121	3	US-10-158-865-1
C 14	662.8	23.0	5966	3	US-08-956-171B-22
C 15	662.8	23.0	5966	3	US-08-781-986A-22
16	630.6	21.9	8729	3	US-09-543-681A-1857
17	630.6	21.9	8729	3	US-09-453-702B-258
18	630.6	21.9	8729	3	US-10-114-170-258
19	630.6	21.9	87563	3	US-09-453-702B-57
20	630.6	21.9	87563	3	US-10-114-170-57
21	623.4	21.6	1716	6	US-08-487-429A-3
22	623.4	21.6	1716	6	PCT-US96-05320A-541
23	617.8	21.4	1878	3	US-09-489-039A-2045
24	612	21.2	3164	3	US-09-710-279-3828

25	612	21.2	3234	3	US-09-710-279-3853	Sequence 3853, Ap
26	578.6	20.1	1716	3	US-09-710-279-117	Sequence 117, App
27	578.6	20.1	1722	3	US-09-134-001C-2189	Sequence 2189, Ap
28	578.4	20.1	1704	3	US-09-328-352-1786	Sequence 1786, Ap
29	572.6	19.9	1701	3	US-09-252-991A-10316	Sequence 10316, A
C 30	557	19.3	1653	3	US-09-252-991A-10546	Sequence 10546, A
31	551.8	19.1	1833	3	US-09-602-777A-13	Sequence 13, Appl
32	528.4	18.3	1625	3	US-09-602-777A-15	Sequence 15, Appl
C 33	512	17.8	1797	3	US-09-601-198-89	Sequence 89, Appl
34	502	17.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
35	502	17.4	4411529	3	US-09-103-840A-1	Sequence 1, Appl
36	424.4	14.7	3475	3	US-09-710-279-4032	Sequence 4032, Ap
37	356.6	12.4	951	3	US-09-252-991A-10513	Sequence 10513, A
C 38	213.2	7.4	636	3	US-09-252-991A-10653	Sequence 10653, A
39	210.2	7.3	684	3	US-10-012-819-153	Sequence 153, App
40	173.2	6.0	363	3	US-10-012-819-159	Sequence 159, App
41	101.6	3.5	613	3	US-09-533-559-7520	Sequence 7520, Ap
42	100.6	3.5	462	3	US-09-543-681A-1913	Sequence 1913, Ap
43	99.2	3.4	312	3	US-09-328-352-1761	Sequence 1761, Ap
44	97.4	3.4	300	6	PCT-US96-05320A-543	Sequence 543, App
C 45	97.4	3.4	309	3	US-09-710-279-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-08-920-095-1
; Sequence 1, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/431,041
; APPLICATION NUMBER:
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-920-095-1

Query Match 37.3%; Score 1075.2; DB 2; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

QY 182 AAAATTTAAACAGGAGTAATAGGTGAAACTCACACCCAAAGACGAAAGTTCTTG 241
Db 78 ACACITTTAAGAAATGAGAGATGAGATGAACTCACCCAAAGAGTTAGATAAGTTGAT 137
QY 242 TTATATATTAGCGGCGAAGTGGCTAGAAAGCGAAAGCAGAGGCTTAAAGCTCAACAA 301
Db 138 CTCACCTACGCTGAGAAATTTGGCTTAAAAACGAAAGAAAGGCATTAAGCTTAACAT 197
QY 302 CCGAAGCCATTTCTTACATTAAGTGGCCCATATTATGGACGAAGCGCCGCTGGAAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTGCCCATATTATGGAAAGCGAGAGCTGGTAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGGAAAAAAGATGAAGTAATGCC 421
Db 258 ACTGCGCTGAATTTGATGCAAGAGGCGCACCTTTTAAAAACGAGATGATGTGATGAT 317
QY 422 GGGGTGGGTAATATGGTTCCGATCTAGGTGTAGAGCCACCTTTCTGATGGTACGAAA 481
Db 318 GGGGTGGCAAGCATGATCCATGAGTGGGTATTGAAGCGATGTTTCTGATGGACTAAA 377
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAACCAAGATGAGCACTTCAAAGCGGCGAAGTGAAA 541
Db 378 CTCGTAACCGTGCATACCCCTATTGAGGCCAATGGTAAATTAGTTCTGGTGGTTG- 434
QY 542 TTTGGTTGCCATTAAGACATCGAGCTCAATGCGAGGCAAGAGTAACCGAACTTTGAGGTT 601
Db 435 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGAAGTT 494
QY 602 ACTAATGAAGGGCTAAATCCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAC 661
Db 495 AAAATGTTTGGCAGCAGACCGGTTCAATCGGCTCACTTCCATTTCTTTGAAGTGAAT 554
QY 662 AAGGCATAAAATTCGATCGTGAAGAACCCCTATGGCAAGCCCTAGATATTTCCCTCTGGC 721
Db 555 AGATGCTTAGACTTTGACAGAGAAACCTTTCGGTAAACGCTTAGACATTTGCGAGCGG 614
QY 722 AACACGCTAGCAATTTGGGCGAGACAAACCCGCAAGTGCAGTTGATTCCTTTGGTGGC 781
Db 615 ACAGCGTAAAGATTTGAGCTGGCGAGAAAGAAATCCCGTAGAAATTGATTGACATTTGGCGGT 674
QY 782 AGTAAAAAGTGTATGCAATGAAGCGCTTTGTGATTAACATCGCGATGAAGCCATAAA 841
Db 675 AACAGAGAAATCTTTGGATTAAACGCAATTTGTTGATAGACAAGCAGACAACGAAGCANA 734
QY 842 CATAAAGCGCTTGACAAAGGCAATCTCACGGATTT----- 877
Db 735 AAAATTGCTTTACACAGACTAAAGACGCTGGTGTTCATGGCGCTAAAGCGATGACAC 794
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA-----ATGAAAAAACAGAAATATGTAAAT 926
Db 795 TATGTAATAAACAAATTAAGGAGTAAGAAATGAATAAGATTTAGCAGAAAAAGAAATATGTTCT 854
QY 927 ACCTAGGACCCAAAGGCGATAAGTGGCTTAGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGGTCTCTACAGGCGATAAAGTGGATTTGGGCGATACAGACTTGTATCGCTGAA 914
QY 987 GTAGAACATCACTATACCACTATGGGGAAGAACTTAAATTTTGGCGGGTAAACATATC 1046
Db 915 GTAGAACATCACTATACCACTTATGGCGAAGAGCTTAAATTCGGTGGCGTAAACCCCTA 974
QY 1047 CGTGAGGTTATGGGTACAGACCAATAGCCCTGATGAAACACCCCTAGATTTTAGTCACT 1106
Db 975 AGAAGGCGATGACCAATCTAACAAACCTTAGCAAGAGAGTTGGATTTAATTTACT 1034
QY 1107 AACGCGATGATATCGACTACCGGGATTTTACAAAGCGCATTTGGGATTTAAAAACGGC 1166
Db 1035 AACGCTTTAATCGTGGATTTACCGGTTATTTATAAGCGGATATTGGTATTAAAGATGGC 1094
QY 1167 AAAATCCATGGCATTTGCAAGGAGGAAACGAAGCATGCAAGATGGCGTAAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAGGCGGTAACAAAGCATGCAAGATGGCGTTAAAAACAAT 1154
QY 1227 ATGTCGTGGGTGTGGGCAAGAGCACCTAGAGGGGAAGGTATGATTTATACCGCTGGG 1286

Db 1155 CTTAGCGTAGGTCTCTACTAGAGCTTATAGCGGTGAAAGTTTGTATCGTAACCGCTGAT 1214
QY 1287 GGAATCGATTTCACACACCCACTTCTTCTCCAAACAAATTTCCCTACCGCTCTAGCCAAAT 1346
Db 1215 GGTATTGACACACATCCACTTCTTATTTACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GGGTTTAAACCAATTTTGGAGGCGCACAGGTCCTGTAGATGCAAGATGCGACTACT 1406
Db 1275 GGTGTAACAAACCATGATTGGTGGTAAACCGGTCCTGCTGATGGCACTTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAATTTGCAACGCAATGTTGCGCGCAGAGAGAGATTTCTATG 1466
Db 1335 ATCACTCCAGGCAAGAAATTTAAATATGATGCTCAGAGCGGCTGAAGAAATATTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAAACAAGTA 1526
Db 1395 AATTTAGGTTTCTTGGCTTAAAGGTAAAGCTTCTAAACGATGCGAGCTTAGCCGATCAAAT 1454
QY 1527 GAAGCGGCGGATTTGGTTTAAATTTGATGAAGACTGGGGCAACAAACCAAGTGGATC 1586
Db 1455 GAAGCGGCTGCATTTGGCTTTAAATTTCAAGAGACTGGGGCACCACTCTCTTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGCTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATACA 1646
Db 1515 AATCATCGGTTAGATGTTGGGCAAAATACGATGTCAAGTCTGCTATCCACACAGACT 1574
QY 1647 GTCAATGAGGCGAGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCAT 1706
Db 1575 TTGATGAGCGGTTGTGTAGAGACACTATGGCTGCTATTGTGAGCGCATATGAC 1634
QY 1707 GCCTAACCAATTTAGGAGCGGTTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGCTGGCGGCGACACGCTCTCTGATATTATTTAAAGTAGCCGAT 1694
QY 1767 GAGCTCAATATTTACCTCTCCACACCCCACTATTCCCTATACCATTAATACGGTT 1826
Db 1695 GAACACAACTTCTTCCGCTTCCACTAAACCCCACTCCCTTTCCCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGAT 1886
Db 1755 GCAGAGCATAAGGACATGCTTATGTTGTGACCACTCTGGATTAAGAGATTAAGAGAT 1814
QY 1887 TTACAATTTTCTCAAAGCGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGTCCAT 1946
Db 1815 GTTCAGTTCTGATTTCAAGGATCCGCCCTCAAAACCATTTGCGGCTGAAGACACTTTTGCAT 1874
QY 1947 GATATGGGTGTATCGGATGACAAAGCTCGGATTTGCGAAGCAATGGGCGGTGACGCGAA 2006
Db 1875 GACATGGGGAATTTTCTCAATCACCAATTTCTGACTCTCAAGCGATGGGCGGTGGGTGAA 1934
QY 2007 GTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAGAAATTTGGTAAAGTTCTCT 2066
Db 1935 GTTATCACTAGACTTTGGCAAAACAGCTGAACAAAACGAAGAAAGATTTGGCGCGTTGAAA 1994
QY 2067 GAAGATGCGAAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATC 2126
Db 1995 GAAGAAAAGCGGATACGACCACTTCAGGATCAACCGCTACTTGTCTTAAATACACCAT 2054
QY 2127 AACCCCGCTTTGACCCAGCGGTGAGGATGATATTCGGCTCTGTGGAAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTTCAGTAGAAAGTGGGCAAGATG 2114
QY 2187 GCCGACTTGGTGTGGAAATCTGCTTTTTCGGCTTAAACCCCAATCTGTATCAAA 2246
Db 2115 GCTGACTTGGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAAA 2174
QY 2247 GCGGTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACCGGCTCTGTGCCCCACTCCCCAA 2306
Db 2175 GCGGATTTATTTGCGTTAAGCCAAATGGGCGATGCGAACGCTTCTATCCCTACCCACAA 2234
QY 2307 CCGGTTTATTAACCGGAAATTTTGGGATCACCGCAAGGCGAAATTTGACACCGAGATC 2366

QY 1287 GGAATCGATTACACACCCACTTCTTCTCCACAACAATTCCTACCGCTCTAGCCAAAT 1346
Db 1215 GGTATTGACACACATCCACTTCAATTCACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGTTTACAAACATGTTTGGAGCGGCACAGGTCTGTAGATGCGACGAATGCGACTACT 1406
Db 1275 GGTGTAAACACCATGATTTGGTGTGGAAACCGGTCTGTCTGATGGCACTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAATTTGCACCGCATGTTGCGCGCAGCAGAGAAGTAGTTTCTATG 1466
Db 1335 ATCACTCCAGGCAGAAATTTAAATGGATGCTCAGAGCGGTGAAGAATAATCTCTATG 1394
QY 1467 AATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAAAGTA 1526
Db 1395 AATTTAGTCTTGGCTAAGGTAACGCTTTTAAACGATCGAGCTTTAGCCGATCAAAAT 1454
QY 1527 GAAGCGGGCGGATGCTTTTAAATGTCATGAAGACTGGGGCACAACACCAAGTGGCATC 1586
Db 1455 GAAGCGGTGCGATTTGGCTTTAAATTCAGAAAGACTGGGGCACCACTCTCTTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGGGACAAATACGATGTGCAAGTCTCGTATCCACACAGACACT 1574
QY 1647 GTCATGAGCGAGGTATGATGATGACACCTTAAATGCAATGAACGGGGCGGCGCATCCAT 1706
Db 1575 TTGAATGAAGCGGTGTTGTAGAAAGACACTATGGCTGCTATTTGTGACGCACTATGCAC 1634
QY 1707 GCTTACCAATTCAGGAGCGGTGGAGGACACTCACTCATGTTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGGTGGCGGGACACGCTCTGATATATTAAAGTAGCGGT 1694
QY 1767 GAGCTCAATATTTACCTCTCCACCACCCCACTATTCCTTATACCATTAATACCGTT 1826
Db 1695 GAACAACAATTTCTCCGCTTCCACTAACCCCACTCCCTTTCACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGACATGCCACACCTGACAAACCGCATCCGCGAGAT 1886
Db 1755 GCAGAGCACATGGACATGCTTATGTTGTGCCACACCTTGGATAAAGCAATTAAGAAGAT 1814
QY 1887 TTACAAATTTCTCAAGCGGTATCCGCGCGCTCTATCCGGCTGAAGATGTCTCCAT 1946
Db 1815 GTTCAGTTCGCTGATTCAGAGATCCGCGCTCAAAACCACTTCGGCTGAAGACACTTTTCAT 1874
QY 1947 GATATGGGTGTATCGGATGACAAAGCTCGGATTCGCAAGCAATGGCGGTGCGAGCGAA 2006
Db 1875 GACATGGGATTTTCTCAATCACAGTTCTGACTCTCAAGGATGGCGGTGGGTGAA 1934
QY 2007 GTGATTTCTCGAACTTGGCAGACTGCGGATAAGAAATAAAAGAAATTTGGTAAGCTTCT 2066
Db 1935 GTTATCACTAGAACTTGGCAACAGCTGACAAAAACAAGAAATTTGGCGGTGAA 1994
QY 2067 GAAGATGGCAAGATACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTATC 2126
Db 1995 GAAGAAAAAGCGGATACGCAACTTCAGGATCAACCGCTACTTGTCTAAATACACCAAT 2054
QY 2127 AACCCCGCTTTGACCCACGCGTGAGCGAGTATATCGCTCTGTGGAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCATCGCTCATGGATTAGCGAGTATGTAGTTTCAGTAGAAGTGGGCAAGTG 2114
QY 2187 GCCGACTTGGTGTGGAAATCTCTCCCTTTTTTGGCGTAAACCCCAAAATCGTGATCAAA 2246
Db 2115 GCTGACTTGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATAGTGTGTTCTCTGAAATGGGCGATTTCTAACCGGTCTGTGCCACTCCCAAA 2306
Db 2175 GCGGATTTCAATGCGTTTAAAGCAAAATGGCGGATCGGAACGCTTCTATCCCTTACCCCAAA 2234
QY 2307 CCGGTTTATTACCGGAAATGTTTGGGCATCACCGCAAGCGAAATTTGACACCGATC 2366
Db 2235 CCGGTTTATTACAGAAATGTTTCGCTCATCATGTTAAAGCTAAATACGATGCAAAATC 2294
QY 2367 ACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCGC 2426

Db 2295 ACTTTTGTGTCTCAAGCGGTATTGACAAAAGGCATTAAAGAAATTAGGACTTTGAAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAAACGAC 2486
Db 2355 CAAGTGTTCGGGTAAAAAAATTCGAGAAATATCACTAAAAAGACATGCAATTCACGAC 2414
QY 2487 AAAACGGCAAAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTGAAGTCAATCTCTGAAACTTACCATGTGTTCTGTTGATGGCAAGAA 2474
QY 2547 TGCACTCTAAACCCACTCGCAAGTCCTCTAGCCAGCGCTACACTTTTCTCTAGG 2604
Db 2475 GTAACTTTCTAACACGACCAATAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 2532

RESULT 3

US-08-467-822-19
; Sequence 19, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno


```
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
US-08-467-822-19

Query Match      36.9%; Score 1062.6; DB 2; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;

QY 192 ACAAGAGTAATAGTGAAACTCACACCCAAAGAGCAAGAAAGTTCCTGTTATATATG 251
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 252 CGGGCGAAGTGGTAGAAGCGCAAGAGCGAGAGGCTTAAAGCTCAACCAACCGAGCCA 311
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 89 CGGGCAGATTGGCAGAAGAACGCTTGGCGGTGGTGAAGCTCAATTAACCGAAGCGG 148
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 312 TTGCTTTACATTAGTGCCTATATTATGACGAAGCGCGCTGGAAAAAACCCTTTGCC 371
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 TCGCGCTCATTAGCGCGCTGTGATGGAAGAGCGCGGTGATGGTAATAAAGGTGGCG 208
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 372 AGCTTATGAAGATGTCATGCATTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTA 431
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 209 ATTTGATCAAGAAGCAGCACTTGGCTTAAAAAGAAAAATGTGATGGACGGGTAGCAA 268
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 432 ATATGTTCCCGATCTAGTGTAGAGCCACCTTCTCTGATGGTAGCAAACTTGTAAC 491
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 269 GCATGATTCATGAAGTGGGGATTGAAGCTTAACCTTCCCGATGGAACCAAGCTTGTAACTA 328
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 492 TGAATTTGGCCCATCGAACCAAGATGAGCACTTCAAGCGGCGGAAGTGAATTTGGTGG 551
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 TCCCACTCCGCTAGAGGTAATGGAATATAGCCCCCGCGAGGCTTCTTAA--AA 385
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 552 ATAAAGACATCAGCTCAATGTCAGGCAAGAAAGTAAACCGAACTTGAGGTTACTAATGAAG 611
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 386 ATGAGGACATTACTATTAAACGCGGCAAGAGCAATTAGCTTGAAGTGAAGAAATAAAG 445
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 612 GGCCTAAATCCTTGATGTGGGTAGGCATTTCCATCTTTTGAAGCTAACAAGGCATAA 671
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 446 GCGATCGTCTGTGAGGTGGGATCACATTTCCACTTCTTCGAAGTGAATAAGCTCTTGG 505
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 672 AATTCGATCGTGAAGAGCTATGCAACGCTAGATATTCCTCTGTCGCAACGCTAC 731
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 506 ACTTCGATCGCAAAAAGCTTTTGCAACGCTAGACATTCGATTCGAAACAGCGGTGC 565
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 732 GCATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAG 791
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 566 GCTTTGMAACCGGGAGGNAAGTGTGGAATCATTTGACATCGGCGGGAATNAGCGCA 825
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 792 TGATTTGGCATGAACCGGCTTGTGAATAACATCGCGGATGAACGCCATAAACAAGCGC 851
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 626 TCTATGGCTTTAATTTCTTTGGTGGATCGCAAGCGCATGCCGATGGTAAAAAACTCGGCT 685
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 852 TTGACAAAGCGAAATCTCAGGATTT----- 877
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 686 TAAAAACGCGCTTAAAGAAAAAGTTTTTGGGTCTGTAAACTGCGGTTGTGAAGCGCACTAAAG 745
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 878 ----ATCAAGTAAGGAGACTCCCATGAAATGAAAGAAACAAAGATATGTAAATACCTACG 933
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 746 ATAACATTAAGGAAAAACCATGAAGAGATTTCAAGAAAGANATATGTTCTATGTATG 805
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 GACCCACCAAGGCGATAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGAAGTAGAAC 993
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 806 GTCCCACTACCGGGGATCGTGTGTAGACTCGGCGACACTGATTTGATCTTTAGAAGTGGAGC 865
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 ATGACTATACACCTATGCGGAAGAACTTAAATTTGGCGCGGTTAAAACTATCCGTGAGG 1053
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 866 ATGATTGCACCACTTATGTGTGAAGAGATCAAAATTTTGGGGCGGTAAAACTATCCGTGATG 925
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 GTATGGGTGAGAGCAATAGCCCTGTATGAAAAACACCTAGATTAGTTCATCACTAAGCGCA 1113
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTTGGTGTCTCAACGCC 985
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1114 TGATTATCGACTACACCGGATTTACAAAGCCGACATTGGGATTAAAAACGGCAAAATCC 1173
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 986 TCATTGTGACTATACGGGCAATTTACAAAGCCGACATTGGATTAAGACGGCAAGATTG 1045
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1174 ATGGCAATGGCAAGGCAAGAAACAAAGACATGCAAGATGCGTAAAGCCCTCATATGGTCG 1233
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1046 CAGGCATTTGGCAAGGCGAGGCAATAAGGACATGCAAGATGGCGTAGATAATAATCTTTGCG 1105
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1234 TGGGTGTGGCACAGAAGCACTAGCAGGGAAGGTATGATTATTACCGCTGGGGAATCG 1293
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1106 TAGGTCTCTGTACAGAGGCTTTGGCAGCTGAGGGCTTGAATGTAAACCGCTGGTGCAATCG 1165
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1294 ATTCACACACCCACTTCTTTCTCCACAACAAATTCCTACCGCTTAGCCCAATGGCGTTA 1353
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1166 ATAGGCAATTCACTTATCTCTCCCAACAATCTCTACTGTCTTTGGCAGCGGGTTA 1225
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1354 CAACCAATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCAAGATGCGACTATCACTC 1413
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1226 CAACCAATGTTTGGAGGAGGCACAGGACCTCGGATGGCAGAAATGCGACCACTCACTC 1285
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1414 CGGCAAAATGGAACCTTGCACCGCATGTTGCGCGCAGCAGAAAGATTTCTATGAATGTGG 1473
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1286 CCGGACGCGCTAATCTAAAAAAGTATGTTGCGTGCGAGCGAAGATAGCCCATGAATCTAG 1345
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATCTGTAGAACAACTGTAAGACGG 1533
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAACCTCTTACGCGATCAGATTGAAGCAG 1405
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1534 GCGCAATTTGTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGCAGTCGATCACT 1593
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1406 GGGCAATTTGTTTAAAAATCCACGAAGACTGGGGAAGCACACCTGCGAGCTATTCCACT 1465
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1594 GCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG 1653
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1466 GCCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCACACCGGATACCTTAACG 1525
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1654 AGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGCGCCCATCCATCCCTACC 1713
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1526 AGCGGGCTGTGTAGAAGACACCTTAGCGGCAATGTCGGGCGCACCATCCATCACTTCC 1585
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1714 ACATTCAGGAGCGGCTGGAGGACACTCACTCATGTATTATCAACATGCGAGCGCAGCTCA 1773
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1586 AACTGGAAGGGCTGGGGTGGACACGCTCAGATGTATCAAAATGGCAGGGGAATTTA 1645
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1774 ATATTCATCCCTCTCCACACCCCACTATTCCTATACCATTAATTAACGTTGCGAGAAC 1833
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1646 ACATTCATCCCGCTCTACTAACCAGCACTTCTTTTCAACCAAAAACACTGAAGCCGAGC 1705
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1834 ACTTAGACATGTGCATGACATGCGCACCACTAGACAAAACGATCCGCGAGGATTTACAT 1893
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1706 ACATGACATGTTAATGGTGTGCGCACCACTTGGATAAAAGTATCAAGGAAGATGTCAGT 1765
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1894 TTCTCAAGCGCTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCATCATATGG 1953
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1766 TTGCCATTCGAGGATTCGCCCCCAAACTATCGGGCTGAAGACCAACTCCATGACATGG 1825
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1954 GTGTGATCGCGATGACAAGCTCGGATTCGAAGCAATGGGGCGTGCAGGGGAAGTATTC 2013
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1826 GGATCTTTTCTATCACCAGCTCCGACTCTCAGGCTATGGGACCGGTAGGCGAGGTATCA 1885
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2014 CTCGAACCTTGGCAGACTGCGGATAAGAAATAAAAGAAATTTGGTGAAGCTTCTCTGAAGATG 2073
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1886 CACGCACTTGGCAGACAGCAGACAAAAACAAAAAGAGTTTGGGCGCTTGAAGAGGAAA 1945
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2074 GCAAGATTAACGATTAATTTCCGCAATTAAGCGCTACATCTCCAATACACTATCAACCCCG 2133
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1946 AAGGGGATTAACGACAACTTCCGCACTCAAGCGCTACATCTCTAAATACCACTCAACCCCG 2005
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2134 CTTTGAACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCCCGCACT 2193
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 2006 GGATCCGCATGGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGCAATATGCCGACC 2065
Qy 2194 TGGTGTGTGGATCCTCGCTTTTTTGGGTAAACCCAAAATCGTGATCAAAAGGGGTA 2253
Db 2066 TCGTGCTTTGGAGTCCGGCTTTCTTTGGCATTAAGCCCAATATGATTTAAGGGCGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATGGGGATTTCAAACGCTGTGCGCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGCGCTCTCTCAATGGGGATGCCAATGCGTCTATTCCACCCCTCAGCCGCTCT 2185
Qy 2314 ATTACCGCGAAATGTTTGGGCATCGCGAAGCGGAAATTTGACACGACATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAACAAATTCGACCAATATCACTTCG 2245
Qy 2374 TTTCAAAGTCGCTATGAAATGGGCTGAAAGAAAGCTGGGCTTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGCTTACAAGCAGGGATCAAGAAGAACTAGGCTAGATCGCGCGCAC 2305
Qy 2434 TACCGTCAAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAACGG 2493
Db 2306 CGCCAGTGAATACTGTGCAATATCACTAAAGAGGACCTCAAATTTCAACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGCAACCCCTGAAACCTATAAGGTGAAAGTGGATGGCAAGAGGTAACT 2425
Qy 2554 CTAAACCCACCTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTGAGCTAGCGCACTTTATAATTTGTTCTAGG 2476

RESULT 4

US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-432-697-19
Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;
Qy 192 ACAAGAGTAAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATG 251
Db 29 ATAAGGAGTTTAGGATGAAACTTAACGCTTAAGAACTAGACAAGTTAATGCTCAATATG 88
Qy 252 CGGGGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCA 311
Db 89 CGGCGAGATTGGCAGAGAAGACGCTTGGCGGTGGTGTGAAACTCAATTACACCGAAGCGG 148
Qy 312 TTGCTTACATTAGTGGCCCATATTATGACGAAGCGCGCCGTGGGAAAAACCCGTTGCC 371
Db 149 TCGGCTCATTAGCGGCGTGTGATGGAAGAGCGCGTGATGTTAATAAAGCGTGGCGG 208
Qy 372 AGCTTATGGAAGAGTGCATGCACCTTTTGTGAAAAAGATGAAGTAATGCCCGGGTGGGTA 431
Db 209 ATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAGAAAAATGTGTGACGCGGTAGCAA 268
Qy 432 ATATGTTCCCGATCTAGGTGTAGAACCCCTTTCTCTGATGGTACGAACTTGTAACTG 491
Db 269 GCATGATTCATGAAGTGGGATTTGAAGCTAACTCCCGATGGAAACCAAGCTTGTAACTA 328
Qy 492 TGAATTGGCCCATCGAACCCAGATGAGCACCTTCAAGCGGGCGGCAAGTGAATTTGGTTGCG 551
Db 329 TCCACACTCCGGTAGAGATAATGCGCAATTTAGCCCGCGGAGGTCTTCTTAAA--AA 385
Qy 552 ATAAAGCATCGAGCTCAATGCGAGCAAGAAAGTAACCGAACTTGAGTTTACTAATGAAG 611
Db 386 ATAGGACATTACTATTAAACGCGGCAAGAACCAATTAGCTTGAAGTGAATAAATAAG 445
Qy 612 GGCTTAAATCCTTCATGTGGGTAGCCATTTCCACTTTTGAAGCTTAACAGGCACCTAA 671
Db 446 GCGATCGTCTGTGAGGTGGGATCACATTTCCACTTCTTCGAAGTGAATTAAGCTCTTGG 505
Qy 672 AATTTCGATCGTGAATAAGCCCTATGCAAAACGCTTAGATATTCCCTCTCGCAACACGCTAC 731
Db 506 ACTTCGATCGCGCAAAAGCTTTTGCAAAACCCCTAGACATTCGATCTCGAACACGCGTGC 565
Qy 732 GCATTGGGGCAGGACAAACCCGCAAAAGTGCAGTTGATTCTCTTGGTGGCAGTAAAAAAG 791
Db 566 GCTTTGAACCCGGGAGGAAAAAGTGTGGAACCTATTGACATCGCGGGAATAAGCGCA 625
Qy 792 TGATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATTAACATAAAGCGC 851
Db 626 TCTATGGCTTTAAATTCTTTGGTGGATCGCAAGCGGATGCGGATGTTAAACCACTCGGT 685
Qy 852 TTGCAAGGCGAAATCTCACGGATT-----
Db 686 TAAACCGCGCTAAAGAAAAAGTTTGGTCTGTAAACCTCGCGTTGTGAAGCAGCTAAAG 745
Qy 878 ----ATCAAGTAAGGAGACTCCCATGAAAAAGAAAAACAAGAAATATGTAATACCTACG 933
Db 746 ATAAACAATAGGAAAAACCATGAAAAAGATTTACGAAAGAAATATGTTCTATGATG 805
Qy 934 GACCCACCAAGGCGATAAAGTGGCTTTAGAGATACCGATCTTTGGGCGAAGTAGAAC 993

Db 806 GTCCCACTACCGGGATCGTGTAGACTCGCGGACACTGATTTGATCTTAGAAGTGAGC 865
Qy 994 ATGACTATACCCTATGGCGAAGAACTTAAATTTGGCGGGGTAAACCTATCCGTGAGG 1053
Db 866 ATGATTGACCACTTATGTTGAAGAGATCAAAATTTGGGGCGGTAAACCTATCCGTGATG 925
Qy 1054 GTATGGTACAGCAATAGCCCTGTATGAAAACACCCCTAGATTTAGTCACTAAACGCGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAAATTAGATTTGGTGCTCACTAACGCC 985
Qy 1114 TGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTAATAAAGCGCAAAATCC 1173
Db 986 TCATTGGGACTATACCGGCAATTTACAAAGCCGACATTTGGGATTAATAAAGCGCAAAATG 1045
Qy 1174 ATGGCAATTTGGCAAGGAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCATATGGTCG 1233
Db 1046 CAGGCATTTGGCAAGGAGGCAATTAAGGACATGCAAGATGGCGTAAATAAATTTTGGC 1105
Qy 1234 TGGGTGTGGCA CAGAAGCA CTAGCAGGGAAGTATGATTTATTAACGCTGGGGGAATCG 1293
Db 1106 TAGGTCCTGTACAGAGGCTTTGGCAGCTGAGGCTTGATTTGAACCGCTGGTGGCATCG 1165
Qy 1294 ATTACACACCACTCTCTTCTCCCAACAAATTCCTACCGCTCTAGCCAAATGGCGTTA 1353
Db 1166 ATACGCAATTTCACTTTATCTCTCCCAACAAATTCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCAATTTGGAGCGGCGCACAGGCTCTGTAGATGCGCAATGCGACTACTATCACTC 1413
Db 1226 CAACCAATTTGGAGGAGGCGCACAGGCTCTGCGATGCGCAATGCGACCACTCACTC 1285
Qy 1414 CCGGCAAAATGGAACTTGCAACCGCATGTGCGCGCAGCAGAGAGTATTTATGAATGTGG 1473
Db 1286 CCGGACGCGCTAATCTAAAGATGATGTGCGTGCAGCGGAAGATAGCCCATGAATCTAG 1345
Qy 1474 GCTTTTGGCAAGGCAATAGCTCTAGCAAAAAAACAATTGTAGAACAAAGTAGAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTCTTTACGAACCTCTTTTACGCGATCAGATTTGAAGCAG 1405
Qy 1534 GCGCAATTTGTTTAAATTCATGAAGACTGGGCGCACACCAACGATGGCATCACT 1593
Db 1406 GCGCAATTTGTTTAAATTCATGAAGACTGGGGAAGCACCCTGAGCTATTTCACT 1465
Qy 1594 GCTTGAGCGTGGCAGATGATGATGATGCAAGTGTGCAAGTGTGATPCCACACCGCATCAAGT 1653
Db 1466 GCCTCAATGTCGCGATGATGATGATGCAAGTGTGCAAGTGTGATPCCACACCGCATCAAGT 1525
Qy 1654 AGCGAGGTATGTAGATGACACCTTAATGCAATGAACGCGGCGCGCATCTCATGCTTACC 1713
Db 1526 AGCGGCGTGTGTAGAAGACACCTTAGAGCGGATTTGCGGGCGCACCATCCATACCTTCC 1585
Qy 1714 ACATTGAGGAGCGGTGGAGGACCTCACTGATGTTATCACCATGGCAGGCGAGCTCA 1773
Db 1586 ACATGAAGGGGTGGGGTGGACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTCTACCTCTCCACACCCCACTATTCCCTATACCTTAATTAATGCTTGCAGAAC 1833
Db 1646 ACATTCTACCGGCTCTCTAATCCCGACCACTTCTTCCACCAAAACACTGAAGCGGAGC 1705
Qy 1834 ACTTACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCACTTGGATAAAGTATCAAGGAAGATGTGCACT 1765
Qy 1894 TTTCTAAAGCGGTATCGCGCGGCTCTATCCGCGCTGAAGTGTGCTCCATGATGG 1953
Db 1766 TTTGCGGATTTGAGGATTTGCGCGGCAAACTATCGCGCTGAAGACCACTCCATGACATGG 1825
Qy 1954 GTGTGATCGGATGACAAGCTGGATTTGCAAGCAATGGGGCTGCGAGCGAGTGATTC 2013
Db 1826 GGAATCTTTCTATCCAGCTCGACTCTCAGGCTATGGACCGTAGGCGAGGTGATCA 1885
Qy 2014 CTCGAATTTGGCAGCTGGGGTAAGAATAAAGAAATTTGGTAGCTTCTCTGAAGATG 2073
Db 1886 CAGGCACTTTGGCAGCAGCAGACAAACAAAGAGATTTGGCGCTTTGAAGAGGAAA 1945

Qy 2074 GCAAAGATAACGATAATTTCCGCATTAAAGCGCTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGCGGATAACGACAACTTCCGCATCAAGCGTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGAAGATCGCGACT 2193
Db 2006 GGAATCGCGCATGGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAAAATACGCGCACC 2065
Qy 2194 TGGTGTGTGGAATCTTCCTTTTGGCGGTAAACCCCAAAATCGTGTATCAAAAGCGGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGCTTTCTTTGGCATTAAGCCCAATATGATTAATTAAGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATCGCGATTTCTAAACGCTCTGTGCCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGGCTCTCTCAATGGCGATGCCAATGCTCTATTCCCACTCAGCCCGTCT 2185
Qy 2314 ATTACCGCAAAATGTTTGGGATCAACGGAAGGGAATTTGACACGACATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAACAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAAAGCTGGCTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGCTTCAAGCGAGGATCAAGAAAGAACTAGGCTAGATCGCGGCGAC 2305
Qy 2434 TACCGGTCAAAAACTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGACAAAAACGG 2493
Db 2306 CGCAGTGAANAATGTCGCAATATCACTAAAGAGACCTCAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACT 2553
Db 2366 CACATATGATGTCAACCTGAAACCTATAAGGTGAAAGTGGATGGCAAGAGGTAACT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 2426 CTAAGCAGCAGATGAATTAGCGCTAGCGCAACTTTATTAATTTGTTCTAGG 2476

RESULT 5

US-08-466-248-19
; Sequence 19, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
;
; US-08-466-248-19
;
; Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
; Best Local Similarity 65.5%; Pred. No. le-263;
; Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;
;
QY 192 ACAAGAGTAATAGGTGAACCTCACACCCCAAGAGCAGCAAGAAAGTTCTTGTATATATG 251
DB 29 ATAAGAGTTTAGGATGAACCTTAACGCCCTAAAGAACTAGACAAAGTTAATGCTCCATATG 89
QY 252 CGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGAGGCTTAAAGCTCAACCAACCCGGAAGCCA 311
DB 89 CGGGCAGATTGGCAGAAAGACGTTTGGCGGTGGTGTGAACCTCAATTACACCGAAGCGG 148
QY 312 TTGCTTACATTAGTGCCTCATTTATGAGCAGAGCGCGCGTGGAAAAAACCCTTGCC 371
DB 149 TCGCGCTCATTAGCGCGGTGATGCAAAAGCGCGTGTATGTTAATAAAGCGTGGCGG 208
QY 372 AGCTTATGAAGAGTGCATGCATTTTGNAAAAAGATGAAGTAATGCCCGGGTGGTA 431
DB 209 ATTTGATGCAAGAGCGAGGACTTGGCTTAAAAAAGAAAAATGTGATGGACGGGTAGCAA 268
QY 432 ATATGGTCCCGATCTAGGTGTAGAAGCCACCTTTCTGATGTTACGAACTTGTAACTG 491
DB 269 GCATGATTCATGAAGTGGGATTTGAAGCTTACCTTCCCGATGGAACCAAGCTTGTAACTA 328
QY 492 TGAATTTGGCCATPCGAACACAGATGAGCATTCAAGCGGGCGGCAAGTGAATTTGGTGG 551
DB 329 TCCACACTCGGTAGAGGATAATGCGCAATTAGCCCGCGGAGTCTTCTTAA--AA 385
QY 552 ATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGTACTAATGAAG 611
DB 386 ATGAGGACATTACTATTAAACCGCGCAAGAGCCATTAGCTTTGAAAGTGAATAAAG 445
QY 612 GGCTAAATCTTGTGATGTGGTAGCATTTCCACTTTTGAAGCTTAAACAGGCACTAA 671
DB 446 GCATCGTCTGTGAGGTGGGATTCATTTCCACTTCTTCGAAGTGAATAAGCTCTTGG 505
QY 672 AATTCGATCGTGAAGAGCCATATGGCAACCGCTAGATATCCCTCTGGCAACACGCTAC 731
DB 506 ACTTCGATCGCGCAAAAGCTTTTGCAAAACGCTAGACATTCGATCTGGAACAGCGGTG 565
QY 732 GCATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAG 791
DB 566 GCTTTGAACCCGGGAGGAAAAAGTGTGAACTCATTTGATGATCGGGCGGAATAAGCGCA 625
;
792 TGATTGGCATGAACGGGCTTTGTGAATAAATCGCGGATGAACGCCCATAAACATAAAGCGC 851
626 TCTATGGCTTTAATTTCTTTGGTGGATCGCAAGCCGATGCGATGTTGTTAAAGAACTCGGCT 685
852 TTGACAAAGCGGAATCTCAGCGATT----- 877
686 TAAACCGCTTAAAGAAAGGTTTGGGTCTGTAAACTCGGTTGTGAAGCGACTAAAG 745
878 ----ATCAAAGTAAGGAGACTCCCATGAATAAAGAAAAAAGAAATATGTAATACCTACG 933
746 ATAAACAATAAGGAAAAACCATGAAGAAAGATTTACGAAAGAAATATGTTCTATGATG 805
934 GACCCACAAAGCGCATAAAGTCGCTTAGAGATACCGATCTTTGGGCGCAAGTAGAAC 993
806 GTCCCACTACCGGGGATCGTGTGTAGACTCGCGCACACTGATTGTGATCTTTAGAAGTGGAGC 865
994 ATGACTATACCACCTATGGCGAAGAACTTAAATTTGGCGGGTAAACATATCTCGTAGG 1053
866 ATGATTGACCACTTATGTGTGAAGATCAAAATTTGGGCGGTAAACATATCTCGTAGG 925
1054 GTATGGGTGAGAGCAATAGCCCTGATGAAACACCCCTAGATTTTAGTCACTCAACGCGA 1113
926 GGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTAGATTTTGGTGTCTCACTAACGCC 985
1114 TGATTATCGACTACACGGGATTTTAAAGCGGACATTTGGGATTTAAACCGGCAAAATCC 1173
986 TCATTGTGCACTATAGCGGCATTTTACAAAGCCGACATTTGGGATTTAAAGACGCGCAAGATTG 1045
1174 ATGGCATTGGCAAGGAGGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCATATGGTCG 1233
1046 CAGGCATTTGGCAAGGAGGCAATAAAGGACATGCAAGATGGCGTAGATTAATCTTTGG 1105
1234 TGGGTGTGGGCACAGAGCACTAGCAGGGAAGGTATGATTATTAACCGCTGGGGAATCG 1293
1106 TAGGTCTGTCTACAGAGCTTTTGGCAGCTGAGGCTTGAATTTAAACCGCTGGTGGCATCG 1165
1294 ATTCAACACCCACTTCTTTTCCCAAAATTTCCCTACCGCTCTAGGCCAATGGCGTTA 1353
1166 ATACGCATATTCATCTTATCTCTCCCAACAAATCTCTACTGCTTTTGGCCAGCGGGTTA 1225
1354 CAACCATTTTGGAGCGGCACACAGTCTGTAGATGGCAAGATGGCACTACTATCACTC 1413
1226 CAACCATGATTGGAGGAGGCAAGGACCTCGGATGGCAAGATGGCACTACTATCACTC 1285
1414 CGGCAAAATCGAACTTGCAACCGCATTTTGGCGCAGCAGAAAGATTTCTATGAATGTGG 1473
1286 CCGGACCGCTAATCTTAAAGATGTTTGGTGCAGCGCAAGATACGCCATGAATCTAG 1345
1474 GCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAAACAGTAGAAGCGG 1533
1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAAACCTCTTTACGGCATCAGATTGAAGCAG 1405
1534 GCGGATTTGTTTAAATTCATGAAGACTGGGCGCACAAACCAAGTCGATCGATCACT 1593
1406 GGGCGATTGTTTAAATTCACGAAAGCTGGGGAAGCACACCTGCACTATTTCCACTC 1465
1594 GCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG 1653
1466 GCCTCAATGTCCCGATGAATACGATGTGCAAGTGGCTTATCCACACCGATACCTTTAAG 1525
1654 AGGCAGGTATGTAGATGACACCCCTAAATGAATGAACGGGCGGCCCATCTCATGCTTACC 1713
1526 AGGCGGCTGTGTAGAAGACACCCCTAGAGCGATTGGCGGGCGCACCATCTCATCTTCC 1585
1714 ACATTGAGGAGCGGGTGGAGGACACTCACCTGATTTATCACCATGGCAGCGAGCTCA 1773
1586 ACCTGAAGGGGCTGGGGGTGGACACGCTCCAGATGTTTATCAAAATGGCAGGGGAATTTA 1645
1774 ATATTCTACCTCTCCACACCCCTATTTCCCTATACCATTAATACGTTTGCAGAAC 1833
1646 ACATTCTACCGCCTCTACTAAACCCGCAATTCCTTTCCAAAAAACACTGAAGCGGAGC 1705
```

```
QY 1834 ACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAAATGGTGTGCCACCACTTGGATAAAAGTATCAAGGAAGATGTGCAGT 1765
QY 1894 TTTCTCAAGCGGTATCCGCGCGGTCTTATCGCGGCTGAAGATGTCTCCATGATGG 1953
Db 1766 TTGCGGATTCGAGGATTTGCGCGCCCAAACTATCGCGCTGAAGACCAACTCCATGACATGG 1825
QY 1954 GTGTGATCCGATGACAAGCTGGATTTCGAAGCAATGGCGGTGCGAGCGAAGTATTC 2013
Db 1826 GATCTTTTCTATCACCAGCTCGGACTCTCAGGCTATGGACGCTAGGCGAGTGATCA 1885
QY 2014 CTCGAATTTGGCAGATCGCGGTAAGAATAAAAAAGAAATTTGTAAGCTTCTTGAAGATG 2073
Db 1886 CAGGCATTTGGCAGACAGCAGACAAAAACAAAAAGAGTTTGGCGCTTGAAGAGGAAA 1945
QY 2074 GCAAGATTAACGATTAATTTCCGATTAAGCGGTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGCGGATAACGACAACTTTCCGATCAAAACGCTACATCTCTAAATACACCATCAACCCCG 2005
QY 2134 CTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGAAGAGGCGAAGATCGCGACT 2193
Db 2006 GGATCGCGATCGGATTTCTGACTATGTGGGCTCTGTGAAGTGGCAATATAGCCGACC 2065
QY 2194 TGGTGTGTGGAATCTCGCTTTTGGCGGTAAAAACCCAAATCTGTATCAAGGCGGTA 2253
Db 2066 TGTGCTTTGGAGTCGGCTTTCTTTGGCATTTAAGCCCAATATGATTTAAGGGCGGAT 2125
QY 2254 TGGTGTCTTCTGAAATGGCGGATTTAAACGCGTCTGTGCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGGCGTCTCTCAAAATGGCGATGCCAATGCGTCTATTCCACCCCTCAGCCCGTCT 2185
QY 2314 ATTACGGGAATGTTGGGATCAGGCAAGCGGCGAATTTGACACCATCATCTTTT 2373
Db 2186 ATTACCGGAAATGTTGGACACCATGGGANAACAAATTCGACACCAATATCACTTTCCG 2245
QY 2374 TTTCAAAGTCGCCATGAAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGCAAGTTTC 2433
Db 2246 TGTCCNAGCGCTTACAGGCGAGGATCAAGAGNACTAGGCTAGATCGCGGCAAC 2305
QY 2434 TACCGGTCAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGG 2493
Db 2306 CGCCAGTGAAGAACTGTGCAATATCACTAAAGAGGACCTCAAAATTCACAGATGTGACCG 2365
QY 2494 GAAAATCACCGTCATCGGAAACCTTCGAGTCTTTGTAGATGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGTAACCCCTGAAACCTATAAGGTGAAAGTGGCAAGAGGTAACCT 2425
QY 2554 CTAACCCACCTCGCAAGTGCCTCTAGCCGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAGCAGCAGATGAATGAGCCTTAGCCCAACTTTATATTGTTCTAGG 2476
```

RESULT 6

```
US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)
; US-09-431-705-1
```

Query Match 36.7%; Score 1058.6; DB 3; Length 4824;

Best Local Similarity 65.7%; Pred. No. 1.4e-262;

Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;

```
QY 197 GAGTAATAGTGTAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATATGCGGGC 256
Db 77 GAGATCTCATGAACTCACCCCAAGAGTTAGATAGTTGATGCTCCACTACGCTGGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCATTGCT 316
Db 137 GAATGGCTAAATAACGCAAGAAAAAGGCATTAAAGCTTAACTATGTAGAAGCAGTAGCT 196
QY 317 TACATTAGTGCCTATATTATGACGAGCGCGCTGGAAAAAAGAAACCGTTGCCAGCTT 376
Db 197 TTGATTAGTGCCTATATTATGGAAGAGCGAGAGCTGGTAAAAAGACTTGGCGCTGAAATTG 256
QY 377 ATGGAAGAGTGCAATGCACATTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAACACAGATGATGATGATGCGGTGGCAAGCATG 316
QY 437 GTTCCCGATCTAGGTGTAGAAGCCACTTTTCTGTATGGTACGAAACTTGTAACTGTGAAT 496
Db 317 ATCCATGAAGTGGTATTGAAGCGATGTTTCTGTATGGGACTTAACTCGTAAACCGTGCAT 376
QY 497 TGGCCCATCGAACCATGAGCACTTCAAGCGCGGGAAGTGAATTTGGTTGCGATAAA 556
Db 377 ACCCTTATTGAGGCCAATGTTAAATTTAGTTCTCTGGTGGTGTG---TTCTTAAAAAATGAA 433
QY 557 GACATCGAGTCAATGTCAGGCAAGAGTAACCCGAACTTCAGGTGTACTAATGAAGGGCT 616
Db 434 GACATCACTATCAACGAGGCAAGAAAGCGGTAGCGTGAAGATTTAAAAATTTGGCGAC 493
QY 617 AAATCCTTGCATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTTAAACAGGCACTAAAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCATTTCTTTTGAAGTGAATAGATGCTTAGACTTT 553
```

Qy	677	GATGTTGAAAAAGCCCTATGGCAAAACGCTAGATATTTCCCTCTGGGCAAACACGCTACGCATT	736
Db	554	GACAGAGAAAAAACTTTTCGGTAAACGCTTAGACATTTCCGAGCGGGACAGCGGTAAAGATT	613
Qy	737	GGGGCAGGACAAACCCGCAAAAGTCGATTGATTTCTCTTGGTGGCAGTAAAAAGTGATT	796
Db	614	GAGCCTGGCGAAGAAAAATCCGTAGAAATGATTTGACATTTGGCGGTAAACAGAAATCTTT	673
Qy	797	GGCATGAAACGGGCTTGTGAATAACATCGCGGATGAAACGCCATAAAAATAAAGCGCTTGAC	856
Db	674	GGATTTTAAACGCAATTGGTTGATAGACAAGCAGACAACGAAAGCAAAAAAATTTGCTTTACAC	733
Qy	857	AAGCGCAATCTCACGGAATTT-----ATCAAGT	884
Db	734	AGAGCTTAAAGAGCGTGTGTTTTCTATGGCGCTTAAAGCGATGACAACTATGTAAAAACAAT	793
Qy	885	AAGGAGACTCCCATGAAA---ATGAAAAACAAGAATATGTAAATACCTACGACGCCACC	941
Db	794	AAGGATAGAAATAGAAAAAGATTAGCAGAAAAAGATATGTTTTCTATGTATGCTCTACT	853
Qy	942	AAAGGGCATAAAGTCGCTTTAGGAGATACCGATCTTTTGGGCAGAAGTAGAACATGACTAT	1001
Db	854	ACAGGCGATAAAGTGAGATTTGGCGGATACAGACTTGATCGCTGAAGTAGAACATGACTAC	913
Qy	1002	ACACCTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGCGTATGGGT	1061
Db	914	ACCATTTATGGCGAAGAGCTTAAATTTGGTGGCGGTAAAAACCCCTAAGAGAAGGCATGAGC	973
Qy	1062	CAGAGCAATAGCCCTGTATGAAAAACACCTAGATTTTAGTCACTACTAACCGGATGATTATC	1121
Db	974	CAATCTAACAACTAGCAAAAGAGTTGGATTTAATTAATCACTAACGCTTTAATCGTG	1033
Qy	1122	GACTACACGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGSCATT	1181
Db	1034	GATTTACACCGGTATTTATAAAGCGGATATTTGGTATTAAAGATGGCAAAATCGCTGGCAATT	1093
Qy	1182	GGCAAGCAGAAACAAGACATGCAAGATGGCGTAAAGCCCTCATATGTTGCTGGGTGTG	1241
Db	1094	GGTAAAGGCGGTAAACAAAGACATGCAAGATGGCGCTTAAAAACAACTCTTAGCTAGGTCT	1153
Qy	1242	GGCACAGAAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGGGGAATCGAATCCACAC	1301
Db	1154	GCTACTGAAGCCTTAGCCGCTGAAGTTTGATCGTAAACGCTGGTGGTATTGACACAC	1213
Qy	1302	ACCCACTTCCTTTTCCACAACAATTCCTACCGCTCTAGCCCAATGCGGTTACAAACATG	1361
Db	1214	ATCCACTTCATTTTCAACCCCAACAAATCCCTACAGCTTTTTCGACGCGGTGTAACACCATG	1273
Qy	1362	TTTGGAGCGGCACAGTCTCTGTAGATGGCAACGNAATCGGACTACTATCACTCCGGCAAA	1421
Db	1274	ATTGGTGGTGGAAACCGGCTCTGCTGTATGGCACTAATGCGACTACTATCACTCCAGGCAGA	1333
Qy	1422	TGGAACCTTGCAACCGCATGTTGGCGCAGCAGAGAAGATTTCTATGAATGTGGGCTTTTGTG	1481
Db	1334	AGAAATTTAAATGGATGCTCAGAGCGCTGAAGAAATTTCTATGAATTTAGGTTTCTTG	1393
Qy	1482	GGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGGCGCAATT	1541
Db	1394	GCTAAAGGTAAACGCTTCTTAAACGATGCGAGCTTTAGCCGATCAAAATGAAGCGGTCGCAATT	1453
Qy	1542	GGTTTTTAAATGTCATGAAGACTGGGCGACACACCAAGTCGATCGATCACTGCTTGAGC	1601
Db	1454	GGCTTTTGCAATTTCAACGAAGACTGGGGCACCACCTCTTCTGCAATCAATCATGCGGTTAGAT	1513
Qy	1602	GTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATAGGCGAGGT	1661
Db	1514	GTTGCGGACAAATACGATGTGCAAGTCGCTATCCCAACAGACACTTTGAATGAAGCCGGT	1573
Qy	1662	TATGTAGATGACACCCCTAAATGCAATGAAACGGCGCGGCATTCATGCTTACCACATTTGAG	1721
Db	1574	TGTTGTAGAAGACACTATGGCTGCTTATGCTGGAACGCACTATGCACTTTTCCACACTGAA	1633

RESULT 7

US-09-431-705-19

US-09-431-705-19
: Sequence 19. Application US/09431705

; sequence 19, Appl
: Patent No. 6585975

; PATENT NO. 6585975
 : GENERAL INFORMATION:

```

; GENERAL INFORMATION:
: APPLICANT: Kleanthous Harold

```

APPLICANT: Kleanthous, Harold

APPLICANT: Londono-Arcilla, Donna
 ADDRESS: Everett, Donna
 CITY: Londono-Arcilla, Donna

APPLICANT: Freeman, Donna
TITLE OF INVENTION: Use of salmonella vectors for

;	TITLE OF INVENTION:	Use of salmonella vectors for	immunization against poliomyelitis infection
;	TITLE OF INVENTION:	Use of salmonella vectors for	immunization against poliomyelitis infection

; TITLE OF INVENTION: vaccination
 THE INVENTOR: 00133/000000

; FILE REFERENCE: 06132/060001

Query Match	Best Local Similarity	Score	1058.6;	DB 3;	Length	4824;
		Mismatches	0;	Indels	38;	Gaps
		Conservative	269;			
197	GAGTAAATAGGTGAACTCACACCCCAAGAGCAAGAAAGTTCTCTTATATATATGCGGC	256				
77	GAGATCTCATGAAACTACCCCAAGAGTTAGATAGTTGATGCTCCACTAGCTGGA	136				
257	GAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAACCGAAGCCATTGCT	316				
137	GAATTGGCTAAAGACGCAAGAAAGGCAATTAAGCTTAACATGATAGAGCAGTAGCT	196				
317	TACATTAGTGGCCATATATATGCAAGAGCGCCCTGGAAAAAAGCCGTTGCCAGCTT	376				
197	TTGATTAGTGGCCATATATATGCAAGAGCGAGAGCTGGTAAAGAGACTGCGGCTGAATTG	256				
377	ATGGAAGATGTCATGCACTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAATATG	436				
257	ATGCAAGAGGGCGCACTCTTTTAAACCAAGATGATGATGGATGGCGTGGCAAGCATG	316				
437	GTTCGCGATCTAGGTGTAAGACCACTTTTCCTGATGGTACGAACTTTGTAACCTGTAAT	496				
317	ATCCATGAGTGGGTATTGAAGCGATGTTCTCTGATGGCATTAACCTGTAACCGTGAT	376				
497	TGGCCCATCGAACAGATGAGCACTTCAAGCGGCGGAAGTGAATTTGGTTGGCATAAA	556				
377	ACCCCTATTGAGGCCAATGGTAAATTAGTTCTCTGGTGAATTG---TTCTTAAAAAATGAA	433				
557	GACATGACCTCAATGCAGGCAAGAGTAACCGAATTTGAGTTTACTTAATGAAGGCGCT	616				
434	GACATCACTATCAACGAAGGCAAAAGACCGGTAGCTGGAAGTTTAAATTTTGGCGAC	493				
617	AAATCTTGCATGTGGTAGCCATTTCCACTTCTTTTGAAGCTTAACGAAGGCACTAAAAATTC	676				
494	AGACGGTTCAATCGGCTCACTTCCATTTCTTTGAAGTGAATAGATGATGATGACTTTT	553				
677	GATCGTGAAGAACCTATGGCAAGCCCTAGATATTCCTCTGGCAACACGCTACGATTT	736				
554	GACAGAGAAAAAATTTGCGTTAAAGCGTTTAGACATTTGGAGCGGACAGCGGTGAATTT	613				

Qy	781	CAGTAA	AAAAAGTGA	ATTGGCA	TGAACGGGCTTGTG	GAATAA	CATCGCGGATGAA	ACGCCATAA	840	
Db	914	TGAACG	CAAAAGTGTATGGTTTAA	ATAATATGTA	ACGA				948	
Qy	841	ACATAA	AGCGCTTGACA	AGCGGAAATCTC	CACGGATTTTAT	CAAGTAA	GAGGACTCC	CAATGA	900	
Db	949	--ATG	GATCAGTTGAAATGGG	AAAGAAATGAGTT					986	
Qy	901	AAATG	AAAAAACAA	AGATATGTAA	ATACCTAC	CGGACCCAC	CAAGGCGATAA	AGTGC	960	
Db	987	CGATG	CTCGAA	AGCAATATCG	GATATGTTTG	EACCAACTGT	CGCGACGCC	CAATCGTT	1046	
Qy	961	TAGAG	ATACCGATCTTT	TGGGCA	AGAGTAGAA	CATGACTAT	ACCACTAT	TGGCGA	GAAC	1020
Db	1047	TGGC	AGATTCAGAA	TTGTTTTAT	CGAAATTGAA	AAAGGACTAT	ACAACTAT	TGAGAT	GAGG	1106
Qy	1021	TTAAAT	TTTGGCGGGT	AAAAACTAT	CCGTGAG	GGTATGGCT	CGAGCAAT			1077
Db	1107	TAAAG	TTTGGCGGCG	CAGGTC	CGAGATGGA	ATGGGG	CGAGCTCTTT	GGCGACAA	1166	
Qy	1078	ATGAAA	CACCC	TAGATTTAGT	CATCACTAAC	CGGATGATTAT	CGACTAC	ACCGGATTT	1137	
Db	1167	GCGAT	GAAATCGGTC	GATCTCGTATTA	ACAAATGCG	ATTATTTG	TGATTA	CACAG	ATTT	1226
Qy	1138	ACAA	CGCGCATTTGGG	ATTA	AAAAACGG	CAAAATCC	ATGGCANTTGG	CAAGCG	AGAAACA	1197
Db	1227	ATAAG	CAGATATCGGC	ATAAAA	AGATGGA	ATGATGCTC	CATAGG	AAAAAGCGGG	GAACC	1286
Qy	1198	AGCAC	ATGCAAGATGG	CGTAAGCCCTCAT	ATGTGCTG	TGGGTGTGGG	CACAGAA	GCAC	TAG	1257
Db	1287	CGTTG	TTAATG	AGCGGGTCG		----ATATG	TGATTTGAG	CAGCA	MACA	1340
Qy	1258	CAGGG	AAGGTATGATTA	TTATTA	ACCGCTGGG	GGAATCGATTC	CACAC	ACCCTCTC	TTTCTC	1317
Db	1341	CCGCA	GAAGGATGATTTGT	CACAGCGG	AGGAATAGAT	GTCTATATTTCA	CTTTATTTG	CC		1400
Qy	1318	CACAA	CAATTCCTTAC	CGCTCTAG	CCAA	TGGCGTTTAC	ACCATGTTTGG	AGCGG	CACAG	1377
Db	1401	CTCAG	CAAAATCGAA	ACCGCTCTTTG	CAATCGG	GTGTGAC	CACTATGATTTGG	CGGAG	CAACAG	1460
Qy	1378	GTCTG	TAGATG	GCACGA	ATGCGACTACTAT	CACTCGG	CGCAATGGA	ACTTGC	CA	1437
Db	1461	GAC	CGGCTAC	GGCA	CAAAATGCC	ACTACTTTGT	ACACGGG	CGCTTGG	AATAT	1520
Qy	1438	TGTTG	CGCGCAG	CAGAGAGTATTT	CTATGA	ATGTGG	CTTTTGGG	CAAAAGG	CAATAGCT	1497
Db	1521	TGCTT	CAAGCAG	CGGAAGAAATCC	CGGATAA	ACTTGGGCTTTTT	AGGAAAGCG	GAAC	CTGTT	1580
Qy	1498	CTAG	CAAAAAAC	AACTTGT	AGACAAGTAG	AAACGGGCGCGAT	TGTTTTAA	ATTG	CATG	1557
Db	1581	CAGAT	GAGGCTCTTT	TAAGG	GAACAAATTC	GAACGGGAGCGG	TGGGATTTAA	AGCTTT	CACG	1640
Qy	1558	AAGAC	TGGGCA	CAACCA	AGTGGCAT	CACTGCTT	GAGCGTGG	CAGATGA	ATACG	1617
Db	1641	AAGAT	TTGGGATCGA	CGGCGGCGCTAT	TGATACAT	GTTTGA	AGTGG	CGGATCG	ATATG	1700
Qy	1618	ATGTG	CAAGTTTGTAT	CCACCG	ATACAGTCA	AGTCAATG	AGGCA	GGTTATGT	AGATGA	1677
Db	1701	ATGTG	CAAGTAG	CGGATTCAT	ACAGACTTT	AAATGA	AGGCGGATTTG	TG	CGAGG	1760
Qy	1678	TAAAT	GCAATGA	AGGGCGGCGCAT	CCATGCCTT	ACCACATTTG	AGGAGCGGGT	TGGAGG	AC	1737
Db	1761	TGAAG	CCATATAG	CGGTG	CGAGTGA	TTCATAC	CTATAC	AGAA	GGGGTGG	1820
Qy	1738	ACTCAC	CTGATGTTAT	CACCATGG	CAGGCTCA	ATATTTCT	ACCTCT	CCCTC	CAAC	1797
Db	1821	ATGCT	CGGATATTA	TAAAG	CGCGGCTTCC	GGATATTTT	GCCTTCTT	CC	CA	1880
Qy	1798	CCA	CTATTCCT	ATACCATTA	ATACGGT	TGCAGAA	CACTTTAG	ACATGCT	CATG	1857
Db	1881	CAACT	CGACCTTAT	CTATCA	TACTTGA	AGAGCATTTAG	ATATGTTA	TGTTA	TGTTG	1940

```

RESULT 9
US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA

```

```

; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,513
; FILING DATE: 11-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/687,645
; FILING DATE: July 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 5783436e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2400
; TYPE: nucleotides
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA
; DESCRIPTION: No
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella aerogenes
; STRAIN: CG253
; INDIVIDUAL ISOLATE:
; CELL TYPE: N/A
; FEATURE:
; NAME/KEY: cDNA encoding mutant urease
; NAME/KEY: 'H219Q
; LOCATION: Modification at position 1312 to
; LOCATION: glutamine
; IDENTIFICATION METHOD: Sequencing
; OTHER INFORMATION:
;
US-08-967-513-1

Query Match      24.1%; Score 693.4; DB 2; Length 2400;
Best Local Similarity 57.1%; Pred. No. 1.7e-168;
Matches 1385; Conservative 0; Mismatches 971; Indels 69; Gaps 4

Qy    195 AGGAGTAATAGGTGAACCTCACCCAAAGACGAGAAAAGTCTTGTTATTATTATCGG 254
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    20  AAGAAGACGTTATGAACTGACCCCCCAGAAAAAACAAGCTGTTCGTTCACGCCG 79
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    255 CGCAAGTGGCTAGAAAGCGCAAAGCAGAGGGCTTAAAGCTCAACCACCCGAAGCATTG 314
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    80  CGTGTGGCGGAGCGTCGCCTGGCCCGCGCTGAAGCTCAACTATCGGAGTCCGTGG 139
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    315 CTTCATTAGTGCCCATATTATGACGAAGCGCGCCGTGGAAAAA AAAAACCGTTGCCACG 374
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    140 CCCTGATCAGCGCCTTTATTATGGAAGCGCTCGGGACGG---CAAAAAGCGTGGCTCGC 196
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    375 TTATGGAAGATGCATGCACCTTTTGA AAAAAGATGAAGTAAATGCCCGGGTGGGTAAATA 434
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    197 TGATGGAGGAAGGCGGTCACTGCTCTGACCCCGCAGACGATGATGAGGGCGTCCCGAAA 256
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    435 TGGTTCCCGATCTAGGTGTAGAAGCCA CTTTCCCTGTAGTGTACGAAA AATTGTAACTGTGA 494
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    257 TGATCCCGGATATCCAGGTGGAAGCACCTTCCCAGACGGCTCGAAGCTGTGTACCGTTC 316
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    495 ATTGGCCCATCGAAACCATGAGCATTCAAAGCGGGCGAAGTGAAA TTTGGTTCCGATA 554
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    317 ACAACCGCATTTATCTGAGTAGCGCATGATCCCGCGTGAATATCACTGTTAAGCCCGTTC 376
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


Qy	375	TTATGGAAGAGTGCATGCACTTTTTCGAAAAAGATGAAATGATGCCCGGGTGGGTAAATA	434
Db	197	 TGATGAGGAAGGCCCGTCACTGCTCACTCCCGCGAGCAGGTGATGGAAGGGCGTCCCGGAAA	256
Qy	435	TGCTTCCCGATCTAGGTGTAGAGCACCCTTTCTGTAGTGGTACGAAACTTTGTAATGTGTGA	494
Db	257	 TGATCCCGGATATCCAGGTGGAAGCACCCTTCCCGACGGCTCGAAGCTGGTCACTGTTTC	316
Qy	495	ATTGGGCCATCGAAACAGATAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGCGATA	554
Db	317	 ACAAACCGATATCTGAGGTAGCGCCATGATCCCCTGGTGAATATCAGTTTAAGCCCGGTTC	376
Qy	555	AAGACATTCGAGCTCAATTCGAGGCAAAAGATGAACGAACTTGTAGGTTTACTAATGAAGGCG	614
Db	377	---AGATAGCCCTGAATACCGGCGGGCAACCTGTCCGTGGTTCGTGTGAGAACCAACGGCG	433
Qy	615	CTAAATCCTTGCATGTGGGTAGCCATTTCCACTTCTTTTGAAGCTACAAGGCACATAAAT	674
Db	434	ATCGCGCGATTCAGGTCCGTTCCACTACATTTCCCGGAGGTAAACCGCGCGCTGAAGT	493
Qy	675	TCGATCGTGAAGAAGCCTATGGCAAAAGCGCTAGATATTCCTCTGTGGCAACACGCTACGCA	734
Db	494	TCGACCGTACGAGCGCGCGCTATCGCTGAATATCCGCGGGGACGCGGTACGCT	553
Qy	735	TTGGGGCAGGACAAAACCCGCAAAAGTCAGTTGATTCCTCTTGTGGTCAGTAAAAAGTGA	794
Db	554	TTGAACCCGGCAGAAACGCGAGGTGCAGTGTGTGGCTTTCGCGGTTCACCGCGCGTCT	613
Qy	795	TTGGCATGAACGGGCTTTGGAATAACATTCGCGATGAACGCCATAAACATAAAGCGCTTG	854
Db	614	TCGGCTTCCGCGCGAGGTCAATGGGCGCTCTGGAGGTAAACGAT-----	657
Qy	855	ACAAGCGAATCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAA	914
Db	658	-----GAGTAAATTTTCAACCCAG	676
Qy	915	GAATATGTAATACTTACCGGACCCACAAAGGCGATAAAGTGCCTTTAGGAGATACCGAT	974
Db	677	 GCCTATGCCGATATGTTTCGCCCCACCGTCGGCGACAGGTGCGCTTGCAGATACCGAG	736
Qy	975	CTTTGGCGAAGTAGAACATGACTATACACTATGGCGGAAGAACTTAAATTTGGCGCG	1034
Db	737	CTGTGGATCGAGTGGAGGACGATTTGACCACTACGGGGAAGAGGTCAAATTCGCGCGC	796
Qy	1035	GGTAAACTATCCGTGAGGTATGGTTCAGACCAATAGCCCTGATGAACACACCTTAGAT	1094
Db	797	GGCAAGTGTATCCGCGACCGCATGGGCGCAGAGGACAGATGCTGGCCCGCACTGTGTGCAC	856
Qy	1095	TTAGTCATCACTAACGGCATGATTTATCGACTACACCGGGATTTTCAAAGCCGACATTTGG	1154
Db	857	CTGGTGTCTACCAACCGCTTATCTGTGATCATCTGGGGATCGTTAAGCCGATATCGGC	916
Qy	1155	ATTAATAACCGCAAAATCCATGGCATTTGGCAAGCGAGGAAACAGGACATGCAAGATGGC	1214
Db	917	GTGAAGGACGCGCGGATCTTCGCCATCGGCAAAAGCGGCAACCCCGACATCCAGCCCAAC	976
Qy	1215	GTAAGCCCTCATATGCTGTGGTGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTT	1274
Db	977	GTCAAC-----ATCCCCCATCGGCGCTGCGACGGGAAGTGAATCGCCGCGGAAGGAAAAAT	1030
Qy	1275	ATTACCGCTGGGGAATCGATTCACACACCACTTCTCTTCTCCACCAAAATTTCCCTACC	1334
Db	1031	GTCAACCGCGCGGGAATCGATACCCATATTTCACTTGTCTGTCCGACAGGCGGGAAGAG	1090
Qy	1335	GCTCTAGCCAATGGCGTTACAAACATGTTTGGAGCGGCGCACAGTCTCTGTAGATGGCAGC	1394
Db	1091	GGCTGTGTCTTGGCGTGAACAATGGTTCGGCGGCGCACCGCCCGCGCGGGGACACC	1150
Qy	1395	AATCGGACTACTATCATTCGCGGCAAAATGAACTTGACCGCATGTTTGGCGCGAGCAGAA	1454
Db	1151	CATGCCACCACTGCACCCCGGGCCGTGGTATATCTCAAGCATGCTGTGACGGGCGCGAC	1210

Qy	1455	GAGTATTCTATGAATGTGGGCTTTTGGCGAAAGCAATAGCTCTAGCAAAAAACAATT	1511
Db	1211	AGCCTGCCGCTCAATATCGCCCTGCTGGCGAAGGGAACGTTTCTCAGCGGATGCCCTG	1270
Qy	1515	GTAGAAACAAGTAGAAGCGGCGGATTTGGTTTAAATTCGATGAAGACTGGGCAACA	1574
Db	1271	CGGAGGAGGTGGCGGAGGGGTTATTGGCGCTGAGATCCAAGAGACTGGGCGGCACC	1330
Qy	1575	CCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGATC	1634
Db	1331	CCGGCGCGATGACTGTGCGTTAACCGTCGCGATGAAATGGAATCCAGGTGCGCCCTG	1390
Qy	1635	CACACGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTAAATGCAATGAACGGG	1694
Db	1391	CACAGGACACCTGATGAATCCGGTTTGTGGAAGACACCCCTGCGCCCATCGCGCGG	1450
Qy	1695	CGCGCATCATCCCTACCAATTTGAGGAGCGGGTGGAGGACACTCACTGATGTATTATC	1754
Db	1451	CGACATCCACACTTCATACCGAAGGGCCGGCGGCGCATGCGCGGACATCATC	1510
Qy	1755	ACCATGGCAGGCGACTCAATATTCTACCTCCTCCACCACCCCACTAATCCCTATACC	1814
Db	1511	ACCGCTCGGCCACCCGACATTTTGGCGTGTGTCACCAACCCAAACGCTGCCCTACACC	1570
Qy	1815	ATTAATACGGTTCGAGAACACTTAGACATGCTCATGACATGCGACACCTAGACAAACGC	1874
Db	1571	CTCAACACCATCGATGAACAATCTCGATATGCTGATGTCTGCCACCATCTGGACCCGGAC	1630
Qy	1875	ATCGGAGGATTTACAAATTTTCTAAACCGGTATCCGCCCGGCTCTATCGCGGCTGAA	1934
Db	1631	ATCGCGAGGACGTGGCTTTTGGCGAGTCGCGATTCGCCGGGAACCATTCGTCGCGAA	1690
Qy	1935	GATGTGCTCATGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGGAACAAATGGGG	1994
Db	1691	GACGTGCTGACGATCTCGCGGCTTCTCGCTCACCTCTCCGATTCGAGGCCATGGGC	1750
Qy	1995	CGTGCAGCGAAGTGAATCTCGAACTTGGCAGACTGGCGGATAAGAAATAAAGAATTT	2054
Db	1751	CGCGTCGGGAAAGTGAATCTTCCGACCTGCGAGTGGCGCATCGCATGAAGGTGCAGCGC	1810
Qy	2055	GGTAAGCTTCTGAAGTAGCAAGATAAGATAATTCGCGATTAAGCGCTACATCTCC	2114
Db	1811	GGAGCGCTGGCGAGGAGACCGGGATAACGACAACTTCGCGGTGAAGCGCTACATCGCC	1870
Qy	2115	AAATACACTATCAACCCCGCTTTGACCACCGGCTGAGCGAGTATATCGGCTCTGTGGAA	2174
Db	1871	AAATACACCATCAACCCGGCGCTGACCCACGGCATCGCACGGAAGTCGATTCATGAG	1930
Qy	2175	GAGGGCAAGATCGCGCACTTGGTGTGTGGAATCTTGCTTTTTTGGCGTAAACCCAAA	2234
Db	1931	GTGGGTAAGCTGGCTGACCTCGTGGTCTGGTCACACAGCTTCTTCGGCGGTGAACCGGC	1990
Qy	2235	ATCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACGGCTGTG	2294
Db	1991	ACCGTGATCAAAAGCGGCATGATCGGCATCGCGCCGATGGCGGATATCAATGCTCTATT	2050
Qy	2295	CCCACTCCCAACCGGTTTATTACC CGAATCTTTGGGCGCATCAACGCAAGGCGAAATTT	2354
Db	2051	CCGACCCCGCAGCGGTGCACTACCGCCCGATGTTTGGCGCGCTGGGCGCGCCCGCAT	2110
Qy	2355	GACACGACATCACTTTTGTTTCCAAAGTCGCTATGAAATGGCGGTGAAGAAAAAGCTG	2414
Db	2111	CACGTCCGCTCACTTCTCTGTGCGAGGCGCGCAGCCAATGGCGTTGCGAGCGGCTG	2170
Qy	2415	GGCTTAGAGCGCCAGTTCTACCGGTCAAAAACTGGCGTTAACTCAACGAAGAAGACTTC	2474
Db	2171	AACCTGCGCAGCGCATCGCGGTGTGAAAGGCTGCGCTACGGTGCAGAAAGCCGACATG	2230
Qy	2475	AAGTTCAACGACAAAAACCGCAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTGTA	2534
Db	2231	GTGCACAACAGTCTGCAAGCCTAACTACCGTCGACGCCAGACCTATGAGGTGCGGGTG	2290
Qy	2535	GATGGCAAACTCTGCACTCTAAACCCCACTCGCAAGTGCCTCTAGGCCAGCGCTACACT	2594

[illegible]

Qy	423	GGGFGGTAATATGGTTCCCGATCTAGGTGTAGAAGCAACTTTCTCTGATGGTACGAAC	482
Db	564660	GTGTGGCGAAATGGTTCATGAAGTCAGATTTGAAGCTACTTTCCCGCATGGCAGCAAC	564601
Qy	483	TTGTAACTGTGAATTTGGCCCATC---GACCAAGATGAGCACTTCAAAGCGGCGAACTG	538
Db	564600	TTGTACCGTGCATAATCCAAATCAGATAACCGWAGGGTGGGCTTTAGCCCAAAATATA	564541
Qy	539	AAATTTTGGTTCGATAAAGACATCGAGCTCAATGCAGGCAAGAAGTAACCGCAACTTGG	598
Db	564540	AATATCAATGTTGGCTTAAGCCACCTTACAAGGAACAAGATATGATCCAGCGCAAT	564481
Qy	599	GTTACTAATGAAGGCCCTAAATCCTTGC--ATGTGGGTAGGCATTTCCACTTCTTTGAAG	656
Db	564480	ACCAATTAGCCGAAGGCGATATTTCTCGCTAATGTGCGCAGAAAAACCGTAAAAATCGAAG	564421
Qy	657	CTAACAGGCACATAAATTCGATCGTGCMAAAGCCTATGGCAACGCT-----AGATA	710
Db	564420	TAAACAATTCAGCGCAGCGGCCAATTCAGTTGGCTCGCATTCATTTTTTTGMAACCA	564361
Qy	711	TTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTC	770
Db	564360	ATAATGCCCTTAAATTTGACCGCACTTTGGCAGCTGGNATCGCTTAAATGTTCCATCTG	564301
Qy	771	CTCTTGTGTCAGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATG	830
Db	564300	GCAATGCGTGGTTTTGAAACCCGGTGAAGTGAAATCAGTGGAAATAGTTGCTTTTGGTG	564241
Qy	831	AAGCCATAACATAAAGCGCTTGACAAAGGGAATCTCAGGATTTATCAAGTAAGG--	888
Db	564240	GTAACCA-AATCAATTTATGTTTCCATTAATCAAAATTGATGCAAAATATAAGTAGGGCA	564182
Qy	889	AGACTCCACTGAAAAAAGAAAAACAAGAAATATGTAATACTCAGCGACCCACCAAGCGG	948
Db	564181	AGATGGCAATTAACAATTTCAAGAGCGCAATATGTAGCAATTATAGTGTCCAAGTTGGG	564122
Qy	949	ATAAGTGCCTTAGGAGATACCGATCTTTGGGCAAGATGAGAACATGACTATACCACCT	1008
Db	564121	ATAAGTGCCTTTAGGCGATACCAATTTATGGGCAACCAATTTGAACAAGATTTATTGACCA	564062
Qy	1009	ATCGGAGAACCTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTCAAGC-	1067
Db	564061	AAGGTGATGAGTGTAAATTTGGTGGCGGTAAAAAGCGTGCGTGATGATGCTCAAGCG	564002
Qy	1068	-----AATAGCCCTGTAGAAAAACCCCTAGATTTTAGTCATCACTAAACCGCATGATTA	1119
Db	564001	GTAGCGCACTCCGCACAATCCGNAATGTTTGGATTTTGTGATTTACCACTGATGATCA	563942
Qy	1120	TCGA---CTACCGCGGATTTACAAAGCCGCACTTTGGGATTTAAAAACGGCAAAATCCATG	1176
Db	563941	TTGATGCTAAATTTAGCAATTTCAAAGCCGATATTGGTATTTCGTGATGGCGGTATTGGG	563882
Qy	1177	GCATTTGGCAAGGCAAGGAAAAACAAGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGG	1236
Db	563881	GTAATTGCAAGCAGGTAAACCTTGACACCACTGGATAACGTCACACCAAAATATGATTAATCG	563822
Qy	1237	GTCTGGGCACAGAACACTAGCAGGGGAAGTATGATTTATTCGCTGGGGGAATTCGATT	1296
Db	563821	GTGCAAGCAGGAAGTTTCAAGCGTGCAATTTAATTTGCNACCGCTGGTGGTATCGATA	563762
Qy	1297	CACACCCCACTTCTTCTCCACAAATTTCCCTACGCTCTAGCCCAATGGCGTTACAA	1356
Db	563761	CCACATTTACTTTATTTGTTCCACAAAGCACAACATGCAATTTGAAAGTGGCGTTACCA	563702
Qy	1357	CCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCAGAATGCGACTATCATCTCGG	1416
Db	563701	CGTAAATTTGGTGTGGAACCTGGCCCTCTCTGATGGTACACGCAACCACTTGTATCCCTG	563642
Qy	1417	GCAATTTGGAACTTGCACCGCATGTTGCGCCGACAGAGAGTATTTCTATGAATGCGCT	1476
Db	563641	GCGCATTTGGTATATGGAACGTATGTTTCAAGCGGCAGAAAGCCTTCCCGGTAAACGTCCGAT	563582
Qy	1477	TTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCG	1536

Qy	423	GGTGGGTAATATGGTTCCGATCTAGGTGTAGAAGCCACTTTCTGTGATGGTACGAAAC	482	Qy	1477	TTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTTGTAGAAACAAGTAGAAGCGGGCG	1536
Db	564660	GTGTGGCGGAATGGTTCATGAAGTCCAGATTGAAGCTACTTCCCGATGGCAGCAAAAC	564601	Db	563581	TTTTTGTAAAGGCAACTGTGTCAACCCCTAGATCTCTCTCGTGTAGCAAAATTTGAAGCGGGTG	563522
Qy	483	TTGTAACTGTGAATTCGGCCCATC-----GAACAGATGAGACACTTCAAGCGGGCGGAAGTG	538	Qy	1537	CGATTGGTTTTAAATTCGATGAAAGACTGGGCAACAACCAAGTGCAGTCGATCACTGCT	1596
Db	564600	TTGTACCGTGCAATAATCCAAATCAGATAAACCGWGGGTGGGCTTTAGCCACAAAAATAA	564541	Db	563521	CATTAGGTTTTAAAAATCCAGAACTGGGTGCAACGCTTGGCGTGTATTGATCTGCT	563462
Qy	539	AAATTTGGTTGGGATAAAGACATCGAGCTCAATTCAGAGGCAAGAGTAACCGAACTTTGAG	598	Qy	1597	TGAGCGTGGCAGATGAATACGATGTGCAAGTTTTGTATCCACCGGATACAGTCAATGAGG	1656
Db	564540	AATATCAATGGTGGCTTAAGCCACCCCTACAAGGAACAAGATATGATCCAGGGCAAT	564481	Db	563461	TAAAAGTAGCAGATGAATGATATTCAGTGGCCATTACACAGACAGCTAAATGAA	563402
Qy	599	GTTACTAATGAAGGGCTTAATCTTTGC--ATGTGGGTAGCCATTTCACATCTCTTTGAAG	656	Qy	1657	CAGGTTATGTAGATGACACCCCTAAATGAATGAAGAAAGGGCGGCCATCATGCTTACCA	1716
Db	564480	ACCAATTAGCGGACCGCCCAATTCAAGTTGGCTCGCATTAACCATTTTTTTTGAAG	564421	Db	563401	GTGGCTTTTTTGAAGACACATGAAGGATGATGAGCAGGATCATTCATATCTTTCCATA	563342
Qy	657	CTAACAGGCATAAATTCGATCGTGTAAGAAAGCCTATGGCAACGCCT-----AGATA	710	Qy	1717	TTGAGGGGCGGTGGAGGACACTCACCTGATGTATCAACCATCGCAGCGGAGCTCAATA	1776
Db	564420	TAAACAAATTCAGCGACCGCCCAATTCAAGTTGGCTCGCATTAACCATTTTTTTGAAG	564361	Db	563341	CGAGGGCGGAGGTGGTGGTGTGTCATGCACCTGACATCATTAAGAGCGGATGTATTCAAACG	563282
Qy	711	TTCCCTCTGGCAACGCTACGATTTGGGCGAGACAAACCCGCAAGTGCAGTTGATTC	770	Qy	1777	TTCTACCTCTCTCCACCAACCCCACTATTCCCTATACCATTAATACGGTTTCGAGAACACT	1836
Db	564360	ATAATGCCCTTAAATTTTGACCGCACCTTTGGCACGTGGATGCGCTTAAATGTTCATCTG	564301	Db	563281	TATTACTGCTTCAACCAACCCGACTCGTCGGTTTACCAAAAAACACCATTTGATGACATT	563222
Qy	771	CTCTTGGTGGCAGTAAAAAGTGATTGGGCATGAACGGGCTTGTGAATAACATCGCGGATG	830	Qy	1837	TAGACATGTCATGACATGCCACCTAGACAAACGCAATCCGGAGGATTTACAAATTTT	1896
Db	564300	GCAATGGCGTGGTTTTGAACCCGGTGAAGTGAATCAGTGGAAATAGTTAGTTTGGTG	564241	Db	563221	TGGATATGTTGATGGTTTGGCATCACCTTAGATAAACGCTGCCGGAAGACGTAGCTTTTG	563162
Qy	831	AACGCCATAAACAATAAGCGCTTGAACAAGCGAAATCTCACGGATTTATCAAGTAAG--	888	Qy	1897	CTCAAAGCCGTATCCGCCCCGGCTCTATCGGGCTGAAGATGTGCTCCATGATATGGGTG	1956
Db	564240	GTAACCA-AATCATTTATGGTTTCCATAATCAAAATTGATGGCAAAATTAAGGTAGGGCA	564182	Db	563161	CCGATAGCCGTATCCGCCCCGAAACCAATTCAGCAGAAAGATATTTTGCATGATATGGGCG	563102
Qy	889	AGACTCCCATGAATAAAGAAACAAGAATAATGTAATACCTACGACCAACCAAGGGCG	948	Qy	1957	TGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCAAGTGTATTCCTC	2016
Db	564181	AGATGCGATTAAACAATTTCAAGAGCGCAATATGTAGCAACTTATGTGTCACACAGTTGGCG	564122	Db	563101	TCCTTCCATTATGATTTAGATTCAGACTCTCAGCGATGGACAGTATTTGGCGAAGTCTGTTTC	563042
Qy	949	ATAAGTGGCTTATGAGATACCGATCTTTGGCGAGAAGTAGAAACATGACTATACCACT	1008	Qy	2017	GAACCTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCTCGAAGATGGCA	2076
Db	564121	ATAAGTCGTTTATGCGGATACCAATTTATGGCAACCAATTTGAACAGATTTATGACCA	564062	Db	563041	GTACATGGCAAACTGCAGATAAGATGAATAATGCAACCTGGTGTAGTAGTAAAGGA-	562983
Qy	1009	ATGGCGAAGACTTAAATTTGGCGCGGTAAAACTATCGTGAGGATTTGGGTGAGGC-	1067	Qy	2077	AGATTAACGATTAATTTCCGCAATTAAGCGCTACATCTCAATATACACTATCAACCCCGCTT	2136
Db	564061	AAGGTGATGATGATAAATTTGGTGGCGGTAAAGCGTGGTGTGATGGCTCAAAGCG	564002	Db	562982	-----AACGATAACTTCCGTATTAAACGATATATCGGGAATATACCACTCAACCCAGCA	562928
Qy	1068	-----AATAGCCCTGATGAATAACACCCCTAGATTAGTATCATCACTAACCGCATTA	1119	Qy	2137	TGACCCACGCGTGAAGGATATATCGGCTCTGTGGGAAGGGCAAGATGCGCCACTTGG	2196
Db	564001	GTACGGCAACTCGGCACAATCCGAATGATTTGGATTTTGTGATTAACCAAGTATGATCA	563942	Db	562927	TTGCACATGGTATTTGGGAGCATATTGGCTGTTAGAGTGGGTAATAATTCAGATATCG	562868
Qy	1120	TCGA---CTACACGGGATTTACAAAGCGGACATTTGGATTAATAACGGCAAAATCCATG	1176	Qy	2197	TGGTGGGAATCTCGCTTTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATGG	2256
Db	563941	TTGATGCTAAATTAGGCAATTAATCAAGCCGATATTTGGATTTCTGTGATGGCGTATTTGGG	563882	Db	562867	TGTTATGGAAACCGATGTTCTTTGGCGTAAACCTGAAAGTGTGATTAATAAAGGCTTTA	562808
Qy	1177	GCATTGGCAAGGCAAGAAACAAGACATGCAAGATGGCGTGAAGCCCTCATATGGTGGTGG	1236	Qy	2257	TGCTCTTCTGTAAATGGGCGATTTCTAACGCGTCTGTGCCACTCTCCCAACCGGTTTAT	2316
Db	563881	GTAATTGGAACAAGAGTTAACCTGACACCATGATACGTCACCAAAATATGATATCG	563822	Db	562807	TTAGCTATGCGAAAAATGGGCGATCCAAATGCTCAATTCCAACACCGCAACTGTATTCT	562748
Qy	1237	GTGTGGGCACAGAAGCTATGACAGGGGAAGTATGATTAATACCGCTGGGGGAATCGATT	1296	Qy	2317	ACCGCAATGTTTGGGCTACAGGCAAGGGAATTTGACACGAGCATCACTTTTGT	2376
Db	563821	GTGCAAGCAAGGAAGTTTCAACCGGTGCATTTAAATTTGCAACCGCTGGTGTATCGATA	563762	Db	562747	ACCGTCAATGTATCGGTGCAAGGCTTAGCAACCGCAAAACAGCAGTATTCCTTTGTTT	562688
Qy	1297	CACACACCACTTCTCTTCTCCACCAAAATTCCTACCGCTCTAGCCATGGGTTACAA	1356	Qy	2377	CCAAAGTCGCTATGAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTCTTAC	2436
Db	563761	CCCACTTCACTTTATTTGTGCAACAACAAGCAACAATGCAATTTGAAGTGGGTTTACCA	563702	Db	562687	CACAAGCCGTGAAAAAGCTGATATTCGTGGAAGTTTCGGTTTACACAAGAAACCATTTG	562628
Qy	1357	CCATGTTTGGAGCGGCACAGGTCTGTAGATGGCAGATGCGACTATCATCACTCCGG	1416	Qy	2437	CGGTCAAAAATCTCCGCTAACATCACCAAGAAAGACTTCAAGTTCAACAGCAAAAACGGCAA	2496
Db	563701	CGTTAATTTGGTGGAACTGGCCCTGCTGATGGTATACACGCAACCACTTTGACCCCTG	563642	Db	562627	CTGTGAAAGGCTCCGCCAACGCTAGGTAAAAAAGATCTGGTTTCATAATGATTAACACCAA	562568
Qy	1417	GCAATTTGAATTTGCAACCGCATTTTGGCGCAGCAGAGAAGTATTTCTATGAATTTGGGCT	1476	Qy	2497	AAATCAACGTCGATCCGAAAACCTTCGAGGCTCTTTGTAGATGGCAAACTCTGCACTCTA	2556
Db	563641	GCAGTGGTATATGGAACGTAATGTTTTCAAGCGCGCAGAAAGCTTTGCCGCTAAACGTCGGAT	563582	Db	562567	ACATTTACTGTTGATGCTGAACGTTATGAAGTTCAGTGGGAGGAGTTAATTTACCTGTG	562508
				Qy	2557	AACCCACCTCGCAAGTGCCTCTAGCCCGCGCTACACTTTCTTCTCTA	2602

Qy	190	ACACAAGAGTAATAGGTGAAATCTCACCCAAAAGAGCAAGAAAGTCTCTGTGTATATTA	249
Db	4028	AATTGGGGGTTTTAAITTTGCAITTTTACACACGAGAGCAACACAAATTAATGATTGTAGT	3969
Qy	250	TGCGGGCGAAGTGGCTAGAAAGCCGAAGCAGAGAGGGCTTAAAGCTCAACCAACCCGAAGC	309
Db	3968	GCGCGCGAAGTTGCAGTGGTGTAAAGACMGTTGGTTGAACTAAATCATCTCGAGGC	3909
Qy	310	CATTGCTTTACATTAGTCCCATATATTGGACGAAGCGCGCGTGGAAAAAAACCGTTGC	369
Db	3908	ATTAGCTTTAATCAGCATGAATATTATTAGAAGGTGCACGCGATCG---TAAGACCGTTGC	3952
Qy	370	CCAGCTATTGGAAGAGTGCAATGCACTTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGG	429
Db	3851	AGAGTTAATGAGTTATGGTAGACAAATTTAAACAAAAGAGATGTCATGATGGTGTGCGA	3792
Qy	430	TAATATGTTTCCCGATCTAGGTGTAGAAGCACCCTTTCCTGATGGTAGGAACTTGTAAAC	489
Db	3791	ACACATGATTACAGATATCGAAATCGAGGCTACGTTCCCGCATGGTACTAAGTTTAATCAC	3732
Qy	490	TGTGAATTTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGCG---GAAGTGAATAATTTGGTT	548
Db	3731	AGTACATCACCTATTGTTTAAAGGAGCGAGTCAATGTATACCAGGAGAAATATTACAA	3672
Qy	549	CGGATAAGACATCGAGCTCAATCGACGGCAAGAAAGTAACCGAACTTTGAGGTTACTTAATG	608
Db	3671	AAAGTACAGAGGTTGAAATTAATAAACCATCATCTGAAACAGTTATTCGAAAGTTGAAATA	3612
Qy	609	AAGGGCCTAAATCCTTTCGATGTGGTAGCCATTTCCACTTCTTTGAAGCTTAAACAGGCAC	668
Db	3611	CAGGAGATGACCAAAATTCAGATGGGCTCACATTTTTCATTTTATGAAGCAAAATGCAGCAT	3552
Qy	669	TAAATTTTCGATTCGTGAAAAAGCCCTATTGGCAAAACGCTAGATATTCCCTCTGGCAACACGC	728
Db	3551	TAGATTTTCGAACGTGAATGGCATATGGAACAACTATTAGATATTCAGCTGGAGCAGCTG	3492
Qy	729	TACGCAATGGGGCAGGCAAAACCGCAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAA	788
Db	3491	TTTCGATTTGAACCTGGGGAATAAAAAAGAGTTTCAATTTAGTATTGAATATGCTGGCAACGTA	3432
Qy	789	AAGTGATTTGGCATGAACCGGCTTTGTGAA---TAACATCGCGGATGAACGCCATAAACATA	845
Db	3431	AAATTTTGTGTTTCGTGGTATGGTCAATGGTCTCATCGATGATGACGTCTCTATCGCC	3372
Qy	846	AAGCGCTTGACAAGCGCAAAATCTCACGGATTATC-----	880
Db	3371	CAACTGATGAATATGATGAATATGCAGTGTATTCGGAGATAACGGTGTGAAAAACGTGA	3312
Qy	881	-----AAGTAAGGAGACTCCCATGAAAAATGAAAAAACAAGAAATATGTAATATC	928
Db	3311	ATAAAAAAGGAGGAAAAAAGATCATGAGCTTTAAAAATGACGCAAAATCAATATACGAGCTT	3252
Qy	929	CTACGGACCCACMAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGAAGT	988
Db	3251	ATACGGTCCCAACTGTTTGAGATTCCTATTCGTTTAGGTGATACGAATCTATTTGCTCAAA	3192
Qy	989	AGACATGACTATACCACTATGCGCAAGAACTTTAAATTTGGCGGGTAAAACTATCCG	1048
Db	3191	AGAAAAAGACTATGCGGTTTATGGTGAAGAGCTACTTTTGGTGGTGGTAAATCTATTAG	3132
Qy	1049	TGAGGGTATGGGTACAGCAATAGCCC-----TGATGAAAAACCCCTAGATTTAGT	1099
Db	3131	AGACGGTATGGCGCAAAATCCTCGTGTAAACACGTGATGACGTGAACGTTGCAGACCTTGT	3072
Qy	1100	CATCATAACCGGATATGATTATTCGACTACACGGGATTTAACAAGCCGACATTTGGGATTA	1159
Db	3071	CATTCTTAATGCGGTTTATATTCGATTAACGATAAAGTGGTTAAAGCTGATATAGGCATTA	3012
Qy	1160	AAACGGCAAAATCCATCGCATTCGCAAGGCAGGAAACAAGGACATGCAACATGCCGTAA	1219
Db	3011	AAATGGTTTATTTTCCGCCATAGTAATGCGGCAACCCAGATATAATGATTAATGT---	2955
Qy	1220	CCCTCATATGGTTCGTGGGTGTGGGCAAGAAGCACTAGCAGGGGAAGGTATGATTATTAC	1279

Db	2954	---	CGACATTTATATAGGTTCAACAACAGATATCATTTGGCGCTGAAGAGTAAAAATCGTCAAC	2898
Qy	1280	CGCTGGGGGAATCGATTTACACACCCACATTCCTTTCTCCACAACAATTCCTTACCCTCT	1339	
Db	2897	TGCTGGTGGTATTTGATACTCATGTTTATTTATCTGAACAGACGAGGTGCGATT	2838	
Qy	1340	AGCCAAATGGCGTTTACAAACCATGTTTGGAGCGCGCACAGGTCCTGTAGATGGCACGAATGC	1399	
Db	2837	AGAAAGTGGTATTACGACTCATATCGGTGGTGTCTGGTCTTCAGAAGGTTCTAAAGC	2778	
Qy	1400	GACTACTATCATTCCGGGCAATGGAACCTTGCACCGCATGTTTCGGCGCAGCAGAAGTA	1459	
Db	2777	AACAACGTGAACCTCCAGGTCCATGGCATATTCATAGAATGTTAGAAAGCTGCCAAGGTTT	2718	
Qy	1460	TTCTATGAATGTGGGCTTTTGGGCAAGCGCAATAGTCTTAGCAAAAAACACTCTGTAGA	1519	
Db	2717	ACGATTAATGTGGTTTACAGGTAAGGACNAGCAACAAATCCAATGTGCATCTATTGA	2658	
Qy	1520	ACAAGTAGAAGCGGGCGCATGTTGTTTAAATTTGCATGAAGACTGGGGCAACAACAAG	1579	
Db	2657	ACAAATCAATGCGGAGCAATTGGATTAAGAGTACATGAAGACTGGGGTGCAACCAATC	2598	
Qy	1580	TGGATCGATCACTGCTTGAGCGTGCAGATGAATACGATGTGCAAGTTTGTATCCACAC	1639	
Db	2597	TGCTTTGAGTCATGCAATATAGATGTTGCTCATGAATTTGATGTTTCAAATTGCAATTCATGC	2538	
Qy	1640	CGATACAGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGC	1699	
Db	2537	AGATACCTTAAATGAAGCAGGATTTATGAAGACACAATGGGTGCTGTTTAAGACCGTGT	2478	
Qy	1700	CATCCATGCCATCCACATTTGAGGAGCGGTGGAGCACACTCACCTGATGTTTATCACCAT	1759	
Db	2477	ACTTCATATGTACCATCTGAAGTCTGCTGGCGGTGATGGCCCTGATTTTAAATTAATC	2418	
Qy	1760	GGCAGGGAGCTCAATTTCTACCTCTCCACCACCCCACTATTTCCCTATACCAATTAA	1819	
Db	2417	CGCTGCAATTTTCAAATATTTTACCTTCATCTCAAAATCCAACTTTCGCTTTATACACATA	2358	
Qy	1820	TACGGTTGCGAGACACTTAGACATGCTCATGACATGCCACCACCTAGACAAAACGATCCG	1879	
Db	2357	TACTGTAGATGAACATTTAGATATGTAATGATTACTCACATTTAAATGGCGCTATTCC	2298	
Qy	1880	CGAGGATTTACAAATTTTCTCAAAGCGTATCGCGCCCGGCTCTATCGCGCTGAAGATGT	1939	
Db	2297	TGAAGATATCGCATTCGAGATTCACGTATTCGTAAAGAAAACGATTCGACGAGAGATGT	2238	
Qy	1940	GCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGC	1999	
Db	2237	TCTGCAAGATATGGGTGTATTCAGTATGATTAAGTTCCGATTCACAAGCAATGGGCGGTGT	2178	
Qy	2000	AGCGGAAGTATTCCTCGAACCTTGGCAGACTCGGATTAAGATAAAAAAGAAATTCGTAA	2059	
Db	2177	AGGTGAAGTAAATTTACACGAAACATGGCAAGTAGCACATCGCATGAAAGAACACGTTGCTC	2118	
Qy	2060	GCTTCTGAAAGTGGCAAGATAACGATAATTTCCGCATTTAAAGCGCTACATCTCCAAATA	2119	
Db	2117	TTTAGATGGTGAATTTGAACATTAATGATTAATCGCATCAACCGTTTATCGCTAAATA	2058	
Qy	2120	CACTATCAACCCCGCTTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGG	2179	
Db	2057	TACAAATTAACCCAGCAATTACATCGTATTTCTGAATATGTAGGATCTCTATCGAGCCGG	1998	
Qy	2180	CAAGATCGCGACTTGGTGGTGGATCCTGCCTTTTGGGTAAAAACCCAAATCGT	2239	
Db	1997	CAAACTAGCTGCAATTTGCTTATGGGACCCAATTTCTTTGGGGTTAAACCTGAAATTAGT	1938	
Qy	2240	GATCAAAAGCGGTATGGTGGTCTTCTCTGAAATGGGCGGATTTCTAAACCGCTGTGCGCCAC	2299	
Db	1937	TGTTAAGGGCGGATTAATTAATCTCTCGCGTAAATGGCGATGCAATGGTTCTATACCTAC	1878	
Qy	2300	TCCCCAACCGGTTTATTAACGGGAATATGTTTGGGCATACGGCAAGCGGAATTTGACAC	2359	

Db 1877 ATCTGAACGATGAAGTACCGTAAATATGATGGTCAATAGCGGCGGAAACCTTACAAGTAC 1818
Qy 2360 CAGCATCACTTTTGTTCAGAGTCGCCCTATGAAAAATGGCGTGAAGAGAAAGCTGGGCTT 2419
Db 1817 GTCAATGACATCGTGTCTAAACTGCTTATGAAATGGTATCAATCGTGCATTAATTT 1758
Qy 2420 AGAGCGCCAGTCTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTT 2479
Db 1757 AAAACGCATGGTGGTCCAGTTAAAAATATATAGACAAATATCTTAAAGCAGATATGAAAA 1698
Qy 2480 CAACGACAAAAAGCGGAAAAATCACCGTGCATCCGAAAAACCTTCGAGGTCTTGTAGATGG 2539
Db 1697 TAAAGTCAACACCTTAAATAGAGCTGTATCCACAAACATATGAAGTATATAGATGG 1638
Qy 2540 CAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTTACATTTCTT 2599
Db 1637 AGAAAAATTAACAAGTAAATGCAGCAACTGAGTTACCATTAACTCAAGATACTTCTTAT 1578
Qy 2600 CTAGGCACAA 2609
Db 1577 CTAGGAGAA 1568

RESULT 15

US-08-781-986A-22/c
; Sequence 22, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-22

Query Match 23.0%; Score 662.8; DB 3; Length 5966;
Best Local Similarity 55.5%; Pred. No. 2.1e-160;
Matches 1437; Conservative 1; Mismatches 1093; Indels 59; Gaps 6;
Qy 70 TRCTATTATTAATTTTAAATAATCTTATATCATATATAATAATAATTAATTAATTAATTA 129
Db 4148 TGCTTCAATATCGGATATGCTAAAAAATAATAATTTTCTTAATTAACATTTGAACCTTGTATA 4089

Qy 130 TTTAAAGTTAATAAAAGTAACGAAATTAGGACTATAATCCCATTCGCTTTAAATTTA 189
Db 4088 TTTATTTGTTTATCACAACTATAGGATTTAATATAAATTTGTTAAATTTGTTCTAATATGA 4029
Qy 190 ACACAAGGAGTAATAGTGAATCTCACACCAAGAGCAAGAAAAGTTCTTTGTTATATTA 249
Db 4028 AATTGGGGTTTTAATTTGCTTTTACAAACGAGAGCAGACAAATTAATGATGTTAGT 3969
Qy 250 TCGGGCGAGTGGCTAGAAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAAGC 309
Db 3968 GGGCGGGAAGTTGCGCTGCTAAAGCAGCGTGTGAAACTTAAATCATCTCTGAGGC 3909
Qy 310 CATTTGCTTACATTTAGTGCCTATTTATGGACGAAGCGCGCTGGAAAAAAGAAACCGTTGC 369
Db 3908 ATTAGCTTTAATCAGCGATGAATTTATTAGAAAGTGCACGCGATGG--TAAAGCCGTTGC 3852
Qy 370 CCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGGTGG 429
Db 3851 AGAGTTAATGAGTTATGGTAGACAAATCTTAACAAAGAGATGTCATGGATGGTGTGA 3792
Qy 430 TAAATATGGTTCCCGATCTAGGTGTAGAAAGCCACTTTCTCTGATGCTAGCAACTTTGTAAC 489
Db 3791 ACACATGATTACAGATATCGAATTCGAGGCTACGTTCCCGATGGTACTAAGTTAATCAC 3732
Qy 490 TGTGAATTTGCCCATCGAACCATGAGCACTTCAAAGCGGC--GAAGTGAATTTGGTT 548
Db 3731 AGTACATCACCTATTGTTTAAAGGAGGCGAGTCAATGATACCGAGGAAATTTATTACAA 3672
Qy 549 GCGATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAACCGAACTTCAGGTTTACTAATG 608
Db 3671 AAAGTACAGAGTTGAATTAATTAACCATCATCTCTGAAACAGTTATCGAGTTGAAATA 3612
Qy 609 AAGGCGCTTAAATCTTTCATGTGGTAGCCATTTCCACTTCTTTTGAAGCTTAACAGGCAC 668
Db 3611 CAGGAGATCGACCAATTCAGTGGGCTCACATTTTTCATTTTATGAAGCAAAATGCAGCAT 3552
Qy 669 TAAAAATTCGATCGTGAAGAAAGCCATATGCAAAAGCCCTAGATATTTCCCTCTGCAACACGC 728
Db 3551 TAGATTTGAAACGCTGAAATGGCATATGGAACCAATTTAGATATTTCCAGCTGGAGCAGCTG 3492
Qy 729 TAGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTCGTCGAGTAAAA 788
Db 3491 TTCGATTTGAACCTTGGGATTAAGAAAGTTCAATTAGTTGAATATCTCTGCAACGTA 3432
Qy 789 AAGTATTGGCATGAACGGGCTTGTGAA---TAAACATCGCGATGAACCCATTAACATA 845
Db 3431 AATTTTGGTTTTCGTGTATGTCATGTCATGTCATGAGTCAGTCGTCATCGCC 3372
Qy 846 AAGCGTTGAACAGGCGAAATCTCAGGATTTATC----- 880
Db 3371 CAACTGATGAATATGATGAATATGAGGTGTTTCGGAGATAACGGTGTGTAACCGTGA 3312
Qy 881 -----AAGTAAAGGAGACTCCCATGAAATGAAGAAACAGAAATATGTAATAC 928
Db 3311 ATAAAAAGGAGGAAAAAGATCATGAGCTTTTAAATGACGCAAAATCAATATACGAGCTT 3252
Qy 929 CTACGGACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAAGT 988
Db 3251 ATACGCTCCAACTGTTGGAGATTCATTCGTTTAGGTGATACGATCTATTGCTCAAT 3192
Qy 989 AGAATGACTATACCATATGCGAAGAACTTAAATTTGCGCGGGTAAACATATCGG 1048
Db 3191 AGAAAAAGACTATGCGGTTTATGTTGGAAGAGCTACTTTTGTGGTGGTAAATCTATTAG 3132
Qy 1049 TGAGGGTATGGTTCAGAGCAATAGCCC-----TCATGAAACACCTAGATTTAGT 1099
Db 3131 AGACGGTATGCGCGAAAAATCCTCGTGAACACGTCATGACGTCGAGCTTTCGAGACTGT 3072
Qy 1100 CATCACTAAACGCGATGATTATCGACTACACCGGGATTTCAAGCGGACATTTGGGATTA 1159
Db 3071 CATTTCTAATGCCGTTATTTATCGATTACGATAAAGTGGTTAAAGCTGATATAGGCATTAA 3012
Qy 1160 AAACGGCAAAATCCATGGCATTTGGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAG 1219

Db	3011	AAATGGTTATATTTCCCATAGGTATATGCCGCAACCCAGATATATAATGATAATGT---	2955
QY	1220	CCCTCATATGGTGGTGGTGGGACAGAACGACTAGCAGGGGAGGTATGATTTATAC	1279
Db	2954	---CGACATATATATAGGTTCAACAACAGATATCATTTGCCGCTGAAGGTHAAATCGTCAC	2898
QY	1280	CGCTGGGGAATGATTCACACACCCACTTCCTTCTCCCAACAATTCCTACCGCTCT	1339
Db	2897	TGCTGGTGATTTGATACTCATGTTTCAATTTTATTAATCCTGGAACAAGCAGAGGTGCGATT	2838
QY	1340	AGCCAATGGGTTACAAACCATGTTTGGAGGCGCACAGTCTCTGTAGATGGACAGATGC	1399
Db	2837	AGAAAGTGATTTACGACTCATATTTGGTGTTGGTACTGTTGCTTCAAGGTTCTAAGC	2778
QY	1400	GACTACTATCACTCCCGGCAAAATGGAACCTTGCAACCCGATGTTGCGCGCAGCAGAAGTA	1459
Db	2777	AGCACTGTATCTCCAGGTCATGGCATATTCATAGAAATGTTAGAAGTGCCTGAAGTTT	2718
QY	1460	TTCTATGAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACCTTGTA	1519
Db	2717	ACCGATTATGTCGGTTTTACAGGTAAGGACAAGCAACAATCCAACCTGCATCTGA	2658
QY	1520	ACAAGTAGAAGCGGCGCATGTTGTTTTAAATTTGCATGAAGACTGGGGCAACAACAAG	1579
Db	2657	ACAAATCAATCGCGAGCAATTTGGATTAAAGGTACATGAAGACTGGGGTGCACACCATC	2598
QY	1580	TGCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACAC	1639
Db	2597	TGCTTTGAGTCATGCAATTAGATGTTGCTGATGAATTTGATGTTCAAAATTCGCAATTA	2538
QY	1640	CGATACAGTCAATGAGCGCAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCG	1699
Db	2537	AGATACTTTAAATGAGCAGGATTTATGGAAGACACAATGGCTGCTGTAAAGACCGTGT	2478
QY	1700	CATTCATGCTCTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACCAT	1759
Db	2477	ACTTCATATGTACCATACTGAAGTGTGCTGGCGGTGCTGCGCTGATTTAAATTAATC	2418
QY	1760	GGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCACTATTTCCCTATACCAATTA	1819
Db	2417	CGCTGCATTTTCAAAATATTTTACCTTCATCTCAAAATCCAACCTTGCCCTTATACA	2358
QY	1820	TACGGTTGCAGAACACTTAGACATGCTCATGATGCGCACACCTAGACAAACGCAATCG	1879
Db	2357	TACTGTAGATGAACATTTAGATATGTTATGATTAATTAATTAATTAATTAATTAATTA	2298
QY	1880	CGAGGATTTACAATTTTCTCAAAGCGGTATCGGCCCGGCTCTATCGCGGCTGAAGATGT	1939
Db	2297	TGAAGATATCGCATTCGCAGATTCACGTATTCGTAAAGAAACGATTCGACGAGAAGTGT	2238
QY	1940	GCTCCATGATATGGGTGTGATCGGATGAAGTTCGGATTCGCAAGCAATGGGGCGTGC	1999
Db	2237	TCTGCAAGATATGGGTGTATTCAGTATGATTTAGTTCCGATTCACAAAGCAATGGGCG	2178
QY	2000	AGCGCAAGTGATTCCTCGAAGTTGCGAGACTCGGATAGAATAAAGAAATTTGGTAA	2059
Db	2177	AGGTGAAGTAATTAACGAAACATGGCAGTAGACATCCCATGAAAGAAACAACGTTGTC	2118
QY	2060	GCTTCTCTGAAGATGGCAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAATA	2119
Db	2117	TTTAGATGGTATTTTGAACATAATGATAATTAATCGCATCAACGTTATATCGCTAATA	2058
QY	2120	CACTATCAACCCCGCTTTGAACCCAGCGGTGAGCAGATATATCGGCTCTGTGGAAGAGG	2179
Db	2057	TACAAATTAACCCAGCAATTAACATGTTATTTCTGAATATGTAGGATCTATCGAGCCGG	1998
QY	2180	CAAGATCCCGCACTTGGTGTGGAAATCCTGCCCTTTTGGGTTAAACCCCAATCGT	2239
Db	1997	CAAACTAGCTGACATTTGCTTATGGGACCAATTTTCTTTGGGTTAAACCTGAATTAGT	1938
QY	2240	GATCAAGCGGTATGGTGTCTCTCTGAAATGGCGGATTTCAACGCGCTGTGCCCCAC	2299

Search completed: November 29, 2005, 07:30:10
Job time : 550 secs

Db	1937	TGTAAGGGCGGATTAATAACTCTGCGGTAATAATGGCGATGCAAAATGGTCTTATACCTAC	1878
QY	2300	TCCCAACCGGTTTATTACCGCAAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACAC	2359
Db	1877	ATCTGAACCGATGNAAGTACCGTAAATGTATGGTCAATACGCGGAAACCTTTACAAGTAC	1818
QY	2360	CAGCATCATTTTGTTCAAAAGTTCGCTATGAAAAATGGCGTGAAAAAGCTGGGCTT	2419
Db	1817	GTCAATGACATTCGTGTCTAAAACCTGCTTATGAAAAATGGTATCAATCGTCATTAAATTT	1758
QY	2420	AGAGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTT	2479
Db	1757	AAAAACGATGTCGTCCAGTTAAAAATAATTAGACAATTTATCTAAAGCAGATATGAAAA	1698
QY	2480	CAACGACAAAACCGCAAAAAATCACCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGG	2539
Db	1697	TACAGTGAACACCTTAATTTAGACGTTGATCCACAACATATGNAAGTATGTAGATGG	1638
QY	2540	CAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTTACGCCAGCGCTACACTTTCTT	2599
Db	1637	AGAAAAAATTTACAAGTAATGCAGCAACTGAGTTACCATTTAACTCAAAGATACTTCTTATT	1578
QY	2600	CTAGGCACAA 2609	
Db	1577	CTAGGAGGAA 1568	

Db 241 GTTATATTATCGGGGGAAGTGGCTAGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCA 300
Qy 301 ACCGAAGCCATTGCTTACATTAGTGCCTCCATATTATGACGAAGCGCGCGTGGAAAAA 360
Db 301 ACCGAAGCCATTGCTTACATTAGTGCCTCCATATTATGACGAAGCGCGCGTGGAAAAA 360
Qy 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAATATGCC 420
Db 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAATATGCC 420
Qy 421 CGGGTGGGTAAATATGTTCCCGATCTPAGTGTAGAAAGCCACCTTTCTCTGATGTCGAA 480
Db 421 CGGGTGGGTAAATATGTTCCCGATCTPAGTGTAGAAAGCCACCTTTCTCTGATGTCGAA 480
Qy 481 ACTTGTAACCTGTGAATTTGGGCCCATGAAACCCAGATGAGCACTTCAAAAGCGGCGGAAGTAA 540
Db 481 ACTTGTAACCTGTGAATTTGGGCCCATGAAACCCAGATGAGCACTTCAAAAGCGGCGGAAGTAA 540
Qy 541 ATTTGGTTGGGATAAAGACATCGAGCTCAATGCAAGGCAAAAGAGTAAACCGAACTTGAGGT 600
Db 541 ATTTGGTTGGGATAAAGACATCGAGCTCAATGCAAGGCAAAAGAGTAAACCGAACTTGAGGT 600
Qy 601 TACTAATGAAGGCGCTAAATCTCTTGCAATGNGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
Db 601 TACTAATGAAGGCGCTAAATCTCTTGCAATGNGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
Qy 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGAAGCGCTAGATATTCCTCTGG 720
Db 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGAAGCGCTAGATATTCCTCTGG 720
Qy 721 CAACAGCTACGATTTGGGGCAGGACAAACCGCAAGTGCAGTGTGATTTCTCTTGGTGG 780
Db 721 CAACAGCTACGATTTGGGGCAGGACAAACCGCAAGTGCAGTGTGATTTCTCTTGGTGG 780
Qy 781 CAGTAAAAAGTCAATGGCATGAACCGGCTTGTGAATAACATCGCGATGAAGCCATAA 840
Db 781 CAGTAAAAAGTCAATGGCATGAACCGGCTTGTGAATAACATCGCGATGAAGCCATAA 840
Qy 841 ACATAAGCGCTTGCAAGGCGAAATCTCAACGATTTATCAAGTAAAGGAGACTCCCATGA 900
Db 841 ACATAAGCGCTTGCAAGGCGAAATCTCAACGATTTATCAAGTAAAGGAGACTCCCATGA 900
Qy 901 AAATGAAAAAACAAGAAATATGTAATACTTAACCTAAGGACCCCAAGGCGATAAAGTGCCT 960
Db 901 AAATGAAAAAACAAGAAATATGTAATACTTAACCTAAGGACCCCAAGGCGATAAAGTGCCT 960
Qy 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
Db 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
Qy 1021 TTTAAATTTGGCGGGTAAACCTATCGTGAGGGTATGGGTACAGACAAATAGCCCTGATG 1080
Db 1021 TTTAAATTTGGCGGGTAAACCTATCGTGAGGGTATGGGTACAGACAAATAGCCCTGATG 1080
Qy 1081 AAAACACCTTAGATTTAGTCACTCAACGCGATGATTTGCACTACACCGGATTTTACA 1140
Db 1081 AAAACACCTTAGATTTAGTCACTCAACGCGATGATTTGCACTACACCGGATTTTACA 1140
Qy 1141 AAGCGCATTTGGATTTAAAAACGGCAAAATTCATGCGATTTGGCAAGCGAGAAACAAAG 1200
Db 1141 AAGCGCATTTGGATTTAAAAACGGCAAAATTCATGCGATTTGGCAAGCGAGAAACAAAG 1200
Qy 1201 ACATGAAGATGGGTAAAGCCCTCATATGTTGCGGGTATGGGTCAGAGACACTAGCAG 1260
Db 1201 ACATGAAGATGGGTAAAGCCCTCATATGTTGCGGGTATGGGTCAGAGACACTAGCAG 1260
Qy 1261 GGGAAAGTATGATTTATACCGCTGGGGGAATTCGATTCACACACCCACTTCTTTCTCCAC 1320
Db 1261 GGGAAAGTATGATTTATACCGCTGGGGGAATTCGATTCACACACCCACTTCTTTCTCCAC 1320
Qy 1321 AACAAATTTCCCTACCGCTCTAGCCAAATGGCGTTACAAACATGTTTGGAGGCGGCAAGTTC 1380
Db 1321 AACAAATTTCCCTACCGCTCTAGCCAAATGGCGTTACAAACATGTTTGGAGGCGGCAAGTTC 1380

Qy 1381 CTGTAGATGCAAGATGCGCACTACTATCACTCCGGGCAAAATGGAACCTTGCAACCCATGT 1440
Db 1381 CTGTAGATGCAAGATGCGCACTACTATCACTCCGGGCAAAATGGAACCTTGCAACCCATGT 1440
Qy 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
Db 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
Qy 1501 GCAAAAAACAACTTGTAGAAACAGTAGAAGCGGCGCGATTTGGTTTTAAATTTGCAATGAAG 1560
Db 1501 GCAAAAAACAACTTGTAGAAACAGTAGAAGCGGCGCGATTTGGTTTTAAATTTGCAATGAAG 1560
Qy 1561 ACTGGGCGACAACACCAAGTGCAGTCACTCTCTTGAGCGTGGCAGATGAATACGATG 1620
Db 1561 ACTGGGCGACAACACCAAGTGCAGTCACTCTCTTGAGCGTGGCAGATGAATACGATG 1620
Qy 1621 TGCAGGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
Db 1621 TGCAGGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
Qy 1681 ATGCAATGAACCGGCGCGCCATCCATGCTACCAATGAGGAGCGGCTGGAGGACACT 1740
Db 1681 ATGCAATGAACCGGCGCGCCATCCATGCTACCAATGAGGAGCGGCTGGAGGACACT 1740
Qy 1741 CACTGATGTTATCACCATGGCAGGAGCTCAATATTTCTACCTCTCCACACCCCTAA 1800
Db 1741 CACTGATGTTATCACCATGGCAGGAGCTCAATATTTCTACCTCTCCACACCCCTAA 1800
Qy 1801 CTATTCCTATACCATTAATAACGTTGCGAGAACCTTAGACATGCTCATGATGCCACC 1860
Db 1801 CTATTCCTATACCATTAATAACGTTGCGAGAACCTTAGACATGCTCATGATGCCACC 1860
Qy 1861 ACCTAGACAAACGATCCGCGAGATTTACAATTTTCTCAAGCGCTATCCGCCCGGCT 1920
Db 1861 ACCTAGACAAACGATCCGCGAGATTTACAATTTTCTCAAGCGCTATCCGCCCGGCT 1920
Qy 1921 CTATCGCGCTGAAAGTGTGCTCATGATATGGGTGTATCGGATGACAAGCTCGGAT 1980
Db 1921 CTATCGCGCTGAAAGTGTGCTCATGATATGGGTGTATCGGATGACAAGCTCGGAT 1980
Qy 1981 CGCAAGCAATGGGCGTGCAGGCAAGTATTCCTCGAACTGGCAGACTGCGGATAAGA 2040
Db 1981 CGCAAGCAATGGGCGTGCAGGCAAGTATTCCTCGAACTGGCAGACTGCGGATAAGA 2040
Qy 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGTGCAGAAAGATAACGATATTTCCGATTA 2100
Db 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGTGCAGAAAGATAACGATATTTCCGATTA 2100
Qy 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGGCTGAGCGAGTATA 2160
Db 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGGCTGAGCGAGTATA 2160
Qy 2161 TCGGCTCTGTGGAAAGAGGCAAGATCGCGGACTTGGTGGTGGAACTCTGCTTTTGTG 2220
Db 2161 TCGGCTCTGTGGAAAGAGGCAAGATCGCGGACTTGGTGGTGGAACTCTGCTTTTGTG 2220
Qy 2221 GCGTAAACCCAAATCGTGAATCAAGGCGGTATGTGCTCTTCTGAAATGGGCGATT 2280
Db 2221 GCGTAAACCCAAATCGTGAATCAAGGCGGTATGTGCTCTTCTGAAATGGGCGATT 2280
Qy 2281 CTAAACGCGCTGTGCGCCACTCCCAACCGGTTTATTAACGGGAAATGTTTGGGCACTACG 2340
Db 2281 CTAAACGCGCTGTGCGCCACTCCCAACCGGTTTATTAACGGGAAATGTTTGGGCACTACG 2340
Qy 2341 GCAAGGCAAAATTTGACACCGATCACTTTTGTGTTTCCAAAGTCGCCCTATGAAAAATGGCG 2400
Db 2341 GCAAGGCAAAATTTGACACCGATCACTTTTGTGTTTCCAAAGTCGCCCTATGAAAAATGGCG 2400
Qy 2401 TGAAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACCTGCCGTAAACATCA 2460
Db 2401 TGAAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACCTGCCGTAAACATCA 2460

```
QY 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCACCGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCACCGTCGATCCGAAACCT 2520
QY 2521 TCGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Db 2521 TCGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
QY 2581 CCAGCGCTACACTTTCTTAGGCAAAATGCCCTTTGGGGGCAAGTTATTATTAGGAA 2640
Db 2581 CCAGCGCTACACTTTCTTAGGCAAAATGCCCTTTGGGGGCAAGTTATTATTAGGAA 2640
QY 2641 TCTTCATCAACGCACTCGCATCGGTCTTGGGTGCGCATCGTGTCTTTAAACCAAC 2700
Db 2641 TCTTCATCAACGCACTCGCATCGGTCTTGGGTGCGCATCGTGTCTTTAAACCAAC 2700
QY 2701 TTTTCATCTTTAAGCAATGCCATTTTAAATTAATTTAAATTTCTTATAATTAATAT 2760
Db 2701 TTTTCATCTTTAAGCAATGCCATTTTAAATTAATTTAAATTTCTTATAATTAATAT 2760
QY 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTGCTATGTGGGG 2820
Db 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTGCTATGTGGGG 2820
QY 2821 TTGTTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
QY 2881 AGG 2883
Db 2881 AGG 2883
```

RESULT 2

```
US-09-904-994B-13
; Sequence 13, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994B-13
```

Query Match 77.3%; Score 2228; DB 3; Length 2452;

Best Local Similarity 94.3%; Pred. No. 0;

Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```
QY 159 AGGACTATATCCCATTCGCTTTAAATTTAAACAGGAGTAATAGGTGAACACTCACAC 218
Db 1 AGGACTATATCCCATTCGCTTTAAATTTAAACAGGAGTAATAGGTGAACACTCACAC 60
QY 219 CCAAGAGCAAGAAAGTTCTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCAAGAGCAAGAAAGTTCTGTTATATTATGCGGGCGAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGGACCAATTCCTTTACATTAAGTCCCATATTATG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCGGACCAATTCCTTTACATTAAGTCCCATATTATG 180
```

```
QY 339 ACGAAGCGCGCCCTGGAAAAAACCCTTCCAGCTTATGGAAGAGTGCATGCACATTTT 398
Db 181 ACGAGCGCGCTCGTGGCAAAAACCCTTCCGAACTTATGGAAGAGTGTATGCACATTTT 240
QY 399 TGNAAAAAGATGAAGTAATGCGCGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAG 458
Db 241 TGAaaaaaaGACGAGGTGATGCCCGGGTGGGAATATGCTCCCTGATTTGGCGGTGGAAG 300
QY 459 CCACCTTTCTGTATGCTGTAACAACTTTGTAATCTGTAATTTGGCCCATCGAACCATGAGC 518
Db 301 CCACTTTCCCGATGGCACCACCAACTCGTAATCTGTAATTTGGCCCATCGAACCTGATGAC 360
QY 519 ACTTCAAGCGCGCAAGTGAATTTGGTTGCGATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTTAAGCGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTGAATCAACCGCAGTA 420
QY 579 AGAAGTAACCGAACCTTGAGGTTACTAATGAAGGGCTTAATCTTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTACCGAACCTAGAAAGTTACTAACGAAGGACCTTAATCTTTGCATGTGGGTAGCC 480
QY 639 ATTTCCACTTTCTTTGAAGCTAACCAAGCCTAAATTCGATCTGTAATAAAGCCCTTATGCA 698
Db 481 ATTTCCACTTTCTTTGAAGCCTAACCAAGCCTTAATTCGATCGGGAATAAAGCCCTTATGCA 540
QY 699 AACGCTTAGATATTCCTCTGCGAACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAG 758
Db 541 AACGCTTAGATATTCCTCTGCGAACACGCTACGCAATTTGGGCGAGGACAAACCCGTAAG 600
QY 759 TGCAGTTGATTCCTTTGGTGGCAGTAAAGTGAATGCGCATGACGGCTTGTGAATA 818
Db 601 TGCAGTTAATTCCTCTTTGGCGGTAGTAAAGTGAATGCGCATGACGGCTTGTGAATA 660
QY 819 ACATCGCGGATGAACCCATATAACATGAAGCGCTTGACAAAGGCGAATCTCACGATTTA 878
Db 661 ATATTCGCGACGAACCGCATATAACAAAGCGCTAGACAAAGCAAAATCTCACGATTTA 720
QY 879 TCAAGTAAGGAGACTCCCATGAAAAACAAAGAAATATGTAATACTCTACGAGCCC 938
Db 721 TCAAGTAAGGAGACTCCCATGAAAAACAAAGAAATATGTAATACTCTACGAGCCC 780
QY 939 ACCAAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 998
Db 781 ACCACAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 840
QY 999 TATACCACCTATGGCGAAGNACTTAAATTTGGCGCGGTAAACTATCCGTGAGGGGTATG 1058
Db 841 TATACCACCTATGGCGAAGNACTTAAATTTGGCGCGGTAAACTATCCGTGAGGGGTATG 900
QY 1059 GGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCACTAACCGCATGATT 1118
Db 901 GGTGAGAGCAATAGCCCGAGATGAAACACCTTAGATTTAGTGATCACCACGCGATGATT 960
QY 1119 ATCGACTACACCGGATTTTCAAAAGCGCATTTGGGATTTAAAAACGGCAAAATCCATGGC 1178
Db 961 ATTGACTACACCGGATTTTCAAAAGCGCATTTGGGATTTAAAAATGGCAAAATCCATGGC 1020
QY 1179 ATTGGCAAGCGAGGAAACAAAGGATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAAGCGAGGAAACAAAGGATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGT 1080
QY 1239 GTGGGCACAGAACACTAGCAGGGGAAGGTATCATTTATTAACCGTGGGGAAATCGATTCA 1298
Db 1081 GTGGGCACAGAACACTAGCAGGGGAAGGTATCATTTATTAACCGTGGGGGAATCGATTCA 1140
QY 1299 CACACCCACTTCTTTCTCCACAACTTCCCTACCGCTCTAGCCATGCGGTTTACAAAC 1358
Db 1141 CACACCCACTTCTCTCTCCACAACTTCCCTACCGCTCTAGCCATGCGGTTTACAAAC 1200
QY 1359 ATGTTTGGAGCGGACAGGTCTGTAGATGGCAGAAATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCACAGGCCCGGTAGTGGACGAATGCGACTACTCACTCCGGGC 1260
QY 1419 AAATGGAACTTGCACCCGCATGTTTGGCGGCGAGAGAGATTTCTATGAATGTGGGCTTT 1478
```


Db 541 CAAACCGTTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAATAAGTGATGGCATGAAC 600
QY 806 GGGCTTTGTAATAACATCGCGATGAACGCATATAACATAAAAGCGTTTGACAAGGGGAAA 865
Db 601 GGGCTTTGTAATAATATTCGGACGAACGCCATATAACACAAAGCACTAGACRAAGCAAAA 660
QY 866 TCTCAGCGATTTTCAAGTAAGAGACTCCCATGAAATGAAATAAACAAGATAATGTAAA 925
Db 661 TCTCAGCGATTTCAATCAAGTAAGAGACTCCCATGAAATGAAATAAACAAGATAATGTAAA 720
QY 926 TACCTACGAGCACCAACAAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA 985
Db 721 CACCTACGAGCACCAACAAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA 780
QY 986 AGTAGAACATGACTATACCACTCTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 781 AGTAGAACATGACTATACCACTCTATGGCGAAGACTCAAATTTGGCGCGGTAAAACTAT 840
QY 1046 CGGTAGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACACCTTAGATTTAGTCATCAC 1105
Db 841 CGGTAGGGTATGGGTACAGAGCAATAGCCAGATGAAAAACACCTTAGATTTAGTCATCAC 900
QY 1106 TAACGCGATGATTCGATCTACACCGGATTTACAAAGCGGACATTTGGGATTAATAACGG 1165
Db 901 CAAACGCGATGATTCGATCTACACCGGATTTATAAAGCGGACATTTGGTATTAAAAAATGG 960
QY 1166 CAAATTCATGGCATTTGGCAAGGACAGAAACAAAGGACATGCAAGTGGCGTAAAGCCCTCA 1225
Db 961 CAAATTCATGGTATTTGGCAAGGCGGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1020
QY 1226 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGG 1285
Db 1021 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGG 1080
QY 1286 GGAATTCGATTCACACACCACTTCCTTTCTTCCACAAATTCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGTTCGATTCGCACACCACTTCCTCTCTCCCAACAAATTCCTACCGCTCTAGCCAA 1140
QY 1346 TGGCGTTACACCATGTTTGGAGGGGACAGCTCTGTAGATGGCAGCAATGCGACTAC 1405
Db 1141 TGGTGTTCACACCATGTTTGGAGGTGGCACAGTCCGGTAGATGGCAGCAATGCGACCAC 1200
QY 1406 TATCACTCCGGGCAATGGAATTCGCACCGCATGTTGCGCGCAGCAGAGAGATTTCTAT 1465
Db 1201 CATCACTCCGGCAATGGAATTCGCACCGCATGTTGCGCGCAGCTGAAGAGTATTTCTAT 1260
QY 1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAAACAATCGTAGAACAGT 1320
QY 1526 AGAAGCGGCGCATTTGGTTTAAATTCATGAAGACTGGGGCAACACCAAGTCCGAT 1585
Db 1321 AGAAGCGGCGCATTTGGTTTAAATTCATGAAGACTGGGGCAACACCAAGTCCGAT 1380
QY 1586 CGATCACTCTTTAGCGGTGCAGATGAATPACGATGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTCTTTAGCGGTGCAGATGAATPACGATGCAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATAGGCGAGGTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCA 1705
Db 1441 GGTCAATAGGCGAGGTATGTAGATGACACCTTAAATGCGATGAACGGCGCGCCATCCA 1500
QY 1706 TGCCTACCAATAGGCGAGGCGGTGGAGGACATCACTACCTGTATGTTATCACCATGGCAGG 1765
Db 1501 TGCCTACCAATAGGCGAGGCGGTGGAGGACATCACTACCTGTATGTTATCACCATGGCAGG 1560
QY 1766 CGAGCTCAATATTTTACCTCTCTCCACACCCCACTATTTCCCTATPACCATTAATACGCT 1825
Db 1561 CGAGCTCAATATTTTACCTCTCTCCACACCCCACTATTTCCCTATPACCATTAATACGCT 1620
QY 1826 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTTAGACAAACGATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTTAGACAAACGATCCGCGAGGA 1680

QY 1886 TTATCAAAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCCCGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAAGCCGTATCCGCCCGGCTCTATTTGCCGTGAAGATGTGCTCCA 1740
QY 1946 TGNATGGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGCGCTGAGGCGA 2005
Db 1741 TGNATTTGGCGTATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGCGCTGAGGCGA 1800
QY 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTTCTAGAACTTTGGCAAACTGCAGACAGAAATAAAAAAGAAATTTGGTAAGCTTCC 1860
QY 2066 TGAAGATGGCAAGATAACGATAATTTCCGATTTAAGCGGTACATCTTCCAAATACATAT 2125
Db 1861 TGAAGATGGTGAGATAATGACAACTTCCGATCAAAACGCTATATCTTCCAAATACACCAT 1920
QY 2126 CAACCCCGCTTTGACCCACCGCGTGAGCGAGTATATCGGCTCTGTGGAAGGCGCAAGAT 2185
Db 1921 TATCCCGCTTTGACCCATGCGGTGAGCGAGTATATCGGCTCTGTGGAAGGCGCAAGAT 1980
QY 2186 CGCCGACTTGGTGTGTGGAATCTCGCTTTTGGCGTAAACCACAAATCGTGATCAA 2245
Db 1981 CGCCGACTTGGTGTGTGGAATCTCGCTTTTGGTGTAAACCACAAATCGTGATCAA 2040
QY 2246 AGCGGTATGGTGTGTCTTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCCCA 2305
Db 2041 AGCGGTATGGTGTGTCTTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCA 2100
QY 2306 ACCGTTTTATTACCGGAAATGTTGGGCAATCAGCGCAAGCGGCAAAATTTGACACAGCAT 2365
Db 2101 GCGGTTTTATTACCGGAAATGTTGGGCAATCAGCGCAAGCGGCAAAATTTGACACAGCAT 2160
QY 2366 CACTTTTGTTCCAAAGTCGCTATGAAATGCGTGAAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCAAAGTCGCTATGAAATGCGGTGAAAGAAAGCTAGGCTTAGAGCG 2220
QY 2426 CCAAGTTCTACCGGTCAAATACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAAGTGTCTACCGGTGAAATACTGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
QY 2486 CAAAACGGCAAAATCACCGTCGATCGGAAACCTTCGAGGTCTTTGTAGATGGCAAACT 2545
Db 2281 CAAAGCGGCGCATPACATCTCGATCTTAAACCTTAAACCTTGTAGATGGCAAACT 2340
QY 2546 CTGCACCTTAACCCACCTCGCAAGTCCCTTAGCCACGCTACACTTTTCTTAGGC 2605
Db 2341 CTGCACCTTAACCCCGCTCTGAAGTGCCTTAGCCCAACGCTACACTTTTCTTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAT 2405

RESULT 4
US-09-904-994B-10
; Sequence 10, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (693) .. (2399)
US-09-904-994B-10

Query Match      74.9%; Score 2158.6; DB 3; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAGTCTCACCCCAAGAGCAGCAAGAAAGTCTCTGTATATATATGCGGCGAAGTGGCT 265
Db 2 GTGAAGTCTCACCCCAAGAGCAGCAAGAAAGTCTCTGTATATATATGCGGCGAAGTGGCT 61

Qy 266 AGAAGCGGCAAGCAGAGGCTTTAAGCTCAACCAACCCGAGCCATTTGCTTACATAGT 325
Db 62 AGAAGCGCAAGCGGAGGCTTTAAGCTCAACCAACCCGAGCCATTTGCTTACATAGT 121

Qy 326 GCCCATATTATGACCAAGCGCGCTGGAAAAAACCCTTCCCGAGCTTATGGAAGAG 385
Db 122 GCCCATATTATGACCAAGCGCGCGCTGGCAAAAGACCGTTGCGGAACCTTATGGAAGAG 181

Qy 386 TGCATGCACCTTTTGA AAAAAGATGAAGTAATGCCCGGGTGGGTAAATATGGTTCCCGAT 445
Db 182 TGTATGCACCTTTTGA AAAAAGACGAGGTGATGCCCGGTGGGGAAATATGTCCTGAT 241

Qy 446 CTAGGTGTAGAGCCACCTTTCTGTATGTGTAACGAAACTTGTAACCTGTGAATTTGCCCATC 505
Db 242 TTAGGGGTGGAAGTACTTTTCCGATGGCACCACCAACTCTGAACCGGTGAATTTGCCCATC 301

Qy 506 GAACCAAGTNGACACTTCAAGCGGCGGAGTGAATAATTTGGTTGCGATAAGACATCGAG 565
Db 302 GAACCCGATGAACACTTCAAGCGGCGGAGTGAATAATTTGGCTGTGATAAGACATTTGAA 361

Qy 566 CTCATATGCAAGCAAGAGTAACCGAACTTGAGGTTACTAATGAAGGGCCATAATCCTTG 625
Db 362 CTCATATGCAAGCAAGAGTAACTGAGGTTACTAATGAAGGGCCATAATCCTTG 421

Qy 626 CATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTTAAAGGCACTTAAATTCGATCGTGAA 685
Db 422 CATGTGGGTAGCCATTTCCACTTTCTTTGAAGCCAAACAAGGCATTTGAAATTCGATCGGAA 481

Qy 686 AAGCCATATGGCAAAAGCCCTAGATATTCCTCTGGGCAACAGCTACGCATGGGGCAGGA 745
Db 482 AAGCCATATGGCAAAAGCCCTAGATATTCCTCTGGGCAACAGCTACGCATTTGGGGCAGGA 541

Qy 746 CAACCCGCAAGTGCAGTTGATTTCTCTTTGGTGGCAGTAAAAAAGTGAATTTGGCATCAAC 805
Db 542 CAACCCGCAAGTGCAGTTGATTTCTCTTTGGGCGCAGTAAAAAAGTGAATTTGGCATCAAC 601

Qy 806 GGGCTTTGTGAATTAACATCGCGGATGAACGCATAAACATAAAGCGCTTGAACAGGCGAAA 865
Db 602 GGGCTTTGTGAATTAATTTGCAGATGAACGCATAAACATAAAGCGCTTGAACAGGCGAAA 661

Qy 866 TCTCAGGATTTATCAAGTAAAGAGACTCCCATGAAATGAAAAACCAAGATATGTAAA 925
Db 662 TCTCAGGATTTATCAATTAAGAGACTCCCATGAAATGAAAAACCAAGATATGTAAA 721

Qy 926 TACCTACGGACCCCAAGCGCAATGAAGTGCCTTAGGAGATACCCGATCTTTGGGCGAGA 985
Db 722 TACCTACGGACCTTACCAAGCGCAATGAAGTGCCTTAGGAGATACCCGATCTTTGGGCGAGA 781

Qy 986 AGTAGAACATGATATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
Db 782 AGTAGAACATGATATACCACTATGCGGAAGAGCTCAATTTGGCGGGGTAAAACTAT 841

Qy 1046 CCGTGAAGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCATCAC 1105
Db 842 CCGTGAAGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCATCAC 901

Qy 1106 TAAGCGGATGATTCGACTACACCGGATTTACAAAGCCGACATTTGGGATTAAGACGG 1165
Db 902 CAACCGGATGATTTAGTACACCGGATTTTCAAAAGCCGACATTTGGCATTAAGATGG 961
```

```
Qy 1166 CAAAATCCATGGCATTTGGCAAGCGCAAGAAACAAGGACATGCAGATCGCGTAAGCCCTCA 1225
Db 962 CAAAATCCATGGCATTTGGCAAGCGCAAGAAACAAGGACATGCAGATCGCGTAAGCCCTCA 1021

Qy 1226 TATGGTCTGGTGTGGGCACAGAACGACATAGCAGGGGAAGGTATGATTTATACCGCTGG 1285
Db 1022 TATGGTCTGGTGTGGGCACAGAACGATTAGCAGGGGAAGGTATGATTTATACCGCTGG 1081

Qy 1286 GGAATATGATTTACACACCCACTTTCTTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA 1345
Db 1082 GGGGATCGATTTACACACCCACTTTCTTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA 1141

Qy 1346 TGGCGTTTACAAACCATTTTGGAGCGGCACAGGTCCTGTAGATGGCAAGTGGGACTAC 1405
Db 1142 TGGCGTTTACAAACCATTTTGGCGTGGCACAGGTCCTGTAGATGGCAAGTGGGACTAC 1201

Qy 1406 TATCACTCCGGGCAAAATGGAACTTGCACCCGATGTTGCGGCGCAGCAGAGAAGTATTTAT 1465
Db 1202 CATCACTCCGGGCAAAATGGAACTTGCACCCGATGTTGCGGCGCAGCTGAAGAGTATTTAT 1261

Qy 1466 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAT 1525
Db 1262 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGTAAAAACAACCTTGTAGAACAAAT 1321

Qy 1526 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAGACTGGGGGCACACACCAAGTGGCAT 1585
Db 1322 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAGACTGGGGGCACAACTTCAAGTGGCAAT 1381

Qy 1586 CGATCACTGCTTGAAGCGTGGCAGATGAATAGTGTGCAAGTGTGCAAGTGTGATTCACACCGATAC 1645
Db 1382 CGATCACTGCTTGAAGCGTGGCAGATGAATAGTGTGCAAGTGTGATTCACACCGATAC 1441

Qy 1646 AGTCAATGAGCGAGGTATGTAGATGACACCTTAAATGCAATGAACGGGCGGCGCATCCA 1705
Db 1442 GGTCAATGAGCGAGGTATGTAGATGACACCTTGAATGCAATGAACGGGCGGCGCATCCA 1501

Qy 1706 TGCTTACCACTTGAAGGGGCGGTGGAGGACACTCACTGATGTTATCAACATGGCGAGG 1765
Db 1502 TGCTTACCACTTGAAGGGGCGGCGGAGGACACTCACTGATGTTATCAACATGGCGAGG 1561

Qy 1766 CGAGTCAATATTTCTACCTCTCCACACACCCCACTATTCCTATATACCATTAATACGGT 1825
Db 1562 CGAGTCAATATTTCTACCTCTCCACACACCCCACTATTCCTATATACCATTAATACGGT 1621

Qy 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCTAGATAAACGCATCCGCGAGGA 1885
Db 1622 TGCAGAACACTTTAGACATGCTCATGACATGCCACCTAGATAAACGCATCCGCGAGGA 1681

Qy 1886 TTTACAAATTTTCTCAAAAGCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCA 1945
Db 1682 TTTACAAATTTTCTCAAAAGCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCA 1741

Qy 1946 TGATATCGGTGTGATCGCGATGCAAGCTCGGATTTCCGCAAGCAATGGGCGGTGCAAGGCGA 2005
Db 1742 TGATATTTGGCGTGTATCGCGATGCAAGCTCGGATTTCCGCAAGCAATGGGCGGTGCAAGG 1801

Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1802 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAAAAGAAATTTGGTAAGCTTCC 1861

Qy 2066 TGAAGATGGCAAAAGATAACGATAATTTCCGCAATTTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATGTGCAGATAACGATAACTTTCCGCTATCAAAAGCTTACATCTCCAAATACACTAT 1921

Qy 2126 CAACCCCGCTTTGACCCACCGCGTGAAGGATATCGGCTCTGTGGAGAGGGCAAGAT 2185
Db 1922 TAACCCCGCTCTAACCCTATGGGGTAAGGAGTATATCGGCTCTGTGGAGAGGGCAAAAT 1981

Qy 2186 CGCGCATTTGTGTGTGGAAATCTGCTTTTGGGGCTAAAAACCCCAAAATCGTGATCAA 2245
Db 1982 CGCTGATTTGTGTGTGGAAATCTGCTTTTGGGGCTAAAAACCCCAAAATCGTGATCAA 2041

Qy 2246 AGGCGGTATGGTGTGTTCTTCTGAAATGGGCGATTTCTAAACGCGTGTGTGCCCACTCCCA 2305
```

```
Db      2042 AGCGGGTATGGTCTTCTGAAATGGCGACTCAACGCGTCCGTGCTACACCTCA 2101
Qy      2306 ACCGGTTTATTACGCGAAATGTTTGGGCATCACGGCAAGCGGCAAAATTTGACACCCAGCAT 2365
Db      2102 GCGGGTTTATTACGCGAAATGTTTGGGCATCACGGCAAGCGGCAAAATTTGACACCCAGCAT 2161
Qy      2366 CACTTTTGTGTTTCCAAAGTCGCCCTATGAAATATGGCGTGAAAGAAAAGCTGGCGTTAGAGCG 2425
Db      2162 CACTTTTGTGTTTCCAAAGTCGCCCTATGAAATATGGCGTGAAAGAAAAGCTAGCGTTAGAGCG 2221
Qy      2426 CCAAGTTCTACCGGTCAAAAGTCGCCGTAACTACCAAGAAAGACTTCAAGTTCAACGA 2485
Db      2222 CAAAGTGTCTACCGGTCAAAAGTCGCCGTAACTACCAAGAAAGACTTCAAAATTTCAACAA 2281
Qy      2486 CAAAACGGCAAAATCACCGTCGATCCGAAAGCTTCCGAGTCTTGTGTAGATGGCAAACT 2545
Db      2282 CAAGCGGCGCATATCACTGTGATCTCTAAACCTTCGAGGTCTTGTGTAGATGGCAAACT 2341
Qy      2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACATTTTCTTAGGC 2605
Db      2342 CTGCACCTCTAAACCCGCTCTGAAGTGCCTCTAGCCAGCGCTACATTTTCTTAGGC 2401
Qy      2606 ACATG 2611
Db      2402 NCAATG 2407

RESULT 5
US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
US-09-904-994B-7

Query Match      67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy      206 GTGAACCTCACACCCAAAGAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db      3 GTGAACCTCACACCCAAAGAGCAGAGGCTTAAAGCTCAATCAACCCGAAAGCCATTGCTTACATTAGT 122
Qy      266 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db      63 AGAAGCGCAAGACGAGGCTTAAAGCTCAATCAACCCGAAAGCCATTGCTTACATTAGT 122
Qy      326 GCCCATATTATGGAACGAGCGCGCGTGGAAGAAAACCGTTGCCAGCTTATGGAAGAG 385
Db      123 GCCCATATTATGGAACGAGCGCGCGTGGAAGAAAACCGTTGCTGAACTTATGGAAGAA 182
Qy      386 TGCATGCACTTTTGAAGAAAAGATGAAGTAATGCCCCGGGGTGGGTAAATATGTTTCCCGAT 445
Db      183 TGTATGCACTTTTGAAGAAAAGATGAGGTGATGCCCGGTGTGGGGAATATGTTCCCTGAT 242
Qy      446 CTAGGTGTAGAACCCACCTTCTCTGATGTCGAAACTTGTAACTGTGAATTGGCCCCATC 505
```

```
Db      243 TTGGCGGTAGAAGCCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAATT 302
Qy      506 GAACCCAGATGAGCACTTTCAAAGCGGGGAAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db      303 GAACCTGATGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTATGAAGACATTCAG 362
Qy      566 CTCAATGACAGCAAAAGATAACCGAACTTGAGGTTACTAATGAAGGCGCTAAATTCCTTG 625
Db      363 CTCAACGTGGGTAAAGGAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATTCCTTG 422
Qy      626 CATGTGGGTAGCCATTTTCCACTTCTTTGAAGCTTAACAGGCACTTAAATTCGATCGTGAA 585
Db      423 CATGTGGGTAGCCATTTTCCACTTCTTTGAACCAACCAAGGCATTTGAATTCGATCGGAA 482
Qy      686 AAAGCCCTATGGCAAAACCCCTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGA 745
Db      483 AAAGCCCTATGGCAAAACCCCTAGATATTTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGA 542
Qy      746 CAAACCCGCAAAAGTGCAGTTGATTTCTTCTGGTGGCAGTAAAAAAGTGAATTTGGCATGAAC 805
Db      543 CAAACCCGTAAGTGCAGTTAAATCCCTCTTGGCGGTAGTAAAAAAGTGAATTTGGCATGAAC 602
Qy      806 GGGCTTGTGAATTAACATCGCGGATGAACGCCATTAACATAAAGCGCTTGACAAGCGCAAA 865
Db      603 GGGCTTGTGAATTAATTTGGCGACGAACGCCCATTAACACAAGGCACTAGACAAGGCAAAA 662
Qy      866 TCTCACGGATTTTATCAAGTAAGGAGACTCCCATGAAATGAAGAAATGAAGAAATATGTAAG 925
Db      663 TCTCACGGATTTTATCAAGTAAGGAGACTCCCATGAAATGAAGAAATGAAGAAATATGTAAG 722
Qy      926 TACTACGGACCCCAACCAAGCGCATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCAGA 985
Db      723 CACTTACGGACCCCAACCAAGCGCATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCAGA 782
Qy      986 AGTAGAACATGACTATACCACTTATGCGGAAGAACTTAAATTTGGCGGGGTAAACCTAT 1045
Db      783 AGTAGAACATGACTATACCACTTATGCGGAAGAGCTCAAAATTTGGCGGGGTAAACCTAT 842
Qy      1046 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGTAGTAAACCAACCTTAGATTTAGTCAAC 1105
Db      843 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGTAGTAAACCAACCTTAGATTTAGTCAAC 902
Qy      1106 TAAACGATGATTTATCGACTACACCGGATTTTCAAAAGCGACATTTGGGATTAATAAACGG 1165
Db      903 CAAACGATGATTTATCGACTACACCGGATTTTAAAGCCGACATTTGGTATTAATAATGG 962
Qy      1166 CAAAATCCATGGCAATTTGGCAAGCGGAGAA CAAGGACATGCAAGATGCGGTAAGCCCTCA 1225
Db      963 CAAAATCCATGGTATTTGGCAAGCGGAGAA CAAGGACATGCAAGATGCGGTAAGCCCTCA 1022
Qy      1226 TATGTCGTGGGTGTTGGGCA CAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1285
Db      1023 TATGTCGTGGGTGTTGGGCA CAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1082
Qy      1286 GGAATTCGATTCACACACCCACTTCCCTTCTCCACAACTTCCCTACCGCTTAGCCAA 1345
Db      1083 GGGATTCGATTCGACACCCACTTCCCTTCTCCACAACTTCCCTACCGCTTAGCCAA 1142
Qy      1346 TGGCGTTACAAACCATTTTGGAGCGGCACAGGTCTCTGTAGATGGCAAGATGCGACTAC 1405
Db      1143 TGGTGTACAAACCATTTTGGAGGTGGCA CAGGTCCGTTAGTGGCAAGATGCGACCAC 1202
Qy      1406 TATCACTCCGGGCAATGGAATTTGCA CCGCATGTTGCGGCGCAGCAGAAAGTATTTCTAT 1465
Db      1203 CATCACTCCGGGCAATGGAATTTGCA CCGCATGTTGCGGCGCAGCAGAAAGTATTTCTAT 1262
Qy      1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACCACTTGTAGAACAGT 1525
Db      1263 GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACCACTTGTAGAACAGT 1322
Qy      1526 AGAAGCGGCGCGATTTGTTTAAATTTGAAGTGAAGCTGGGCGCAACACCAAGTGGCAT 1585
```

Db 1323 AGAAGCGGGCGGCGATTGGCTTTAAATTGCGTGAAGACTGGGGCACAACACCAAGTGGCAT 1382
QY CGATCACTGCTTGAGCGGTGGCAGATGAATACAGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db CGATCACTGCTTGAGCGGTGGCAGATGAATACAGATGTGCAAGTTTGTATCCACACCGATAC 1442
QY AGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA 1705
Db GGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA 1502
QY TGCTTACCACATTAGGAGCGGTGGAGGACACTCACCTGATGTTATACCATGCGAGG 1765
Db TGCTTACCACATTAGGAGCGGTGGAGGACACTCACCTGATGTTATACCATGCGAGG 1562
QY CGAGCTCAATATTTCTACCTCTCCACACCCCTCACTATTCCTTATACCAATTAATACGGT 1825
Db CGAGCTCAATATTTCTACCTCTCCACACCCCTCACTATTCCTTATACCAATTAATACGGT 1622
QY TGAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACCGATCCGCGAGGA 1885
Db TGAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACCGATCCGCGAGGA 1682
QY TTTACAAATTTTCTCAAGCGGTATCGCCCGGCTCTATCGCGCTGAAGATGTCTCCA 1945
Db TTTACAAATTTTCTCAAGCGGTATCGCCCGGCTCTATCGCGCTGAAGATGTCTCCA 1742
QY TGATATGGGTGTGATCGCATGACAAGCTCGGATTCGCAAGCAATGGGGGTGCGAGCGGA 2005
Db TGATATGGGTGTGATCGCATGACAAGCTCGGATTCGCAAGCAATGGGGGTGCGAGCGGA 1802
QY AGTGATTCCTCGAATCTGGCAGACTGCGGATAGAAATTAAGAAATTTGGTAAGCTTCC 2065
Db AGTGATTCCTCGAATCTGGCAGACTGCGGATAGAAATTAAGAAATTTGGTAAGCTTCC 1862
QY TGAAGATGGCAAGATTAACATTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db TGAAGATGGCGAGATTAACATTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 1922
QY CAACCCCGCTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db TAATCCGCTTTGACCATGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 1982
QY CGCGCACTTGCTGGTGGTGAATCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
Db CGCGCACTTGCTGGTGGTGAATCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2042
QY AGCGGTATGGTGTCTTCTGTAATGGGCGATTTCTAAACGCTGTGCGCACTCCCA 2305
Db AGTGGCATGGTGTCTTCTGTAATGGGCGATTTCTAAACGCTGTGCGCACTCCCA 2102
QY ACCGGTTTATTACCGCAATGTTTGGGCGATCAGGCAAGCGAAATTTGACACCGCAT 2365
Db ACCGGTTTATTACCGCAATGTTTGGGCGATCAGGCAAGCGAAATTTGACACCGCAT 2162
QY CACTTTTGTCTTCCAAAGTCG 2385
Db CACTTTTGTCTCTCAAGCG 2182

RESULT 6

US-10-639-273-1
; Sequence 1, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaglan
; TITLE OF INVENTION: HELICOBACTER BIZZOERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639, 273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337

; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8407
; TYPE: DNA
; ORGANISM: Helicobacter bizzoeronii
US-10-639-273-1

Query Match 41.1%; Score 1184.4; DB 7; Length 8407;
Best Local Similarity 68.5%; Pred. No. 7.3e-260;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;

QY 183 AAATTAAACAAGAGTAATAGTGAAATCAACCCAAAGAGCAAGAAAGTTCTTGT 242
Db 2319 AAATTGGTAGAGGAGTTTAGGATGAAATTAACCCCTAAAGAGCTGGACAAGCTCATGT 2378
QY 243 TATATATGCGGGGAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAGCTCAACCAAC 302
Db 2379 TGCAATTAATGCGGGCGAATGGCTAAAAAAGCGAAGCAAAATGGCGTTAAAGCTAAATATA 2438
QY 303 CCGAAGCATTGTCTTACATTAGTGCCTATATTATGGACGAAGCGCGCTGGAAAAA 362
Db 2439 CTGAGGAGTAGCCCTCATCAGTGCCTATGTATGGAGAAAGCCGCTGCAGGTAAAAA 2498
QY 363 CCGTGGCCAGCTTATCGAAGAGTGCATGCACATTTTGAATAAAGATGAAGTAATGCCG 422
Db 2499 GTGTGGCGGATTTGATCGAAGAGGAGGAGGACATCTCTTAAAGCTGATGATGTCTGCCG 2558
QY 423 GGGTGGTAAATATGGTTCCTGATCTAGTGTAGAGCCACCTTCTTCTGATGGTAGCAAC 482
Db 2559 GTGTAGCCCATATGATCACAAGTGGGGATTGAAGCTAACTTCCCTGATGGGCAAAAC 2618
QY 483 TTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGCGAAGTAAAT 542
Db 2619 TGGTAACCATCCATACCCCGTTGAAGATGGTGGGCATAAATGGCTCCGGTGAAGTGA 2678
QY 543 TTGGTTGCGATAAAGACATCGAGCTCAATCGAGCAAGAAAGTAACCGCAACTTTGAGTTA 602
Db 2679 TTTTGAATAAAGCAAGACATCACTTTGAATGAGGCAAAAGCAACCACTTTAGAAGTGC 2738
QY 603 CTAATGAAGGCGCTAAATCTTGTGATGTGGTAGCCATTTCCACTTCTTTGAAGTAACA 662
Db 2739 ATAAACAAAGCGATCGCCCGCTGCAAGTGGCTCCCACTTCCACTTCTTTGAAGTGAATA 2798
QY 663 AGGCACATAAATTCGATCGTGAAGAGCTATGCAACGCTAGATATTCCTCTGGCA 722
Db 2799 AGCTTTTGGAAATTTGATCGTGAATAAGCCTATGGCAACGCTAGACATTCCTTCTGNA 2858
QY 723 ACAGCTACGATTTGGGGCAGGCAAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCA 782
Db 2859 CCGCTGTGGCTTTGAACCCGCTGAGAAAAAACCCTGGAAATTTGATTCAAATTTGGCGTA 2918
QY 783 GTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGGGATGAACGCCATAAAC 842
Db 2919 ACCAACGCAATTTACGGCTTTAACTCTCTTGTGGATCGCAAGCGCATCTGATGGCAAAA 2978
QY 843 ATAAAGCGCTTGACAAAGCGGAATCTCACGGATTT----- 877
Db 2979 AACTTGTCTCAACCGCGCCAAAGAACATGCTTTTGGTGTGTAATTTGGGTTCCGATA 3038
QY 878 ----ATCAAGTAAGGAGACTCCCATGAAAAATGAAAAA---ACAAGATATGTAATACTCT 930
Db 3039 AAAATAAGGAAGAGCAATCCGATGAAAAAATCTCTCGAAGAAAGATATGTTTCTATGT 3098
QY 931 ACGACCCACCAAGGCGGATAAAGTGGCTTTAGAGATACCGATCTTTGGGCGAGAGTAG 990
Db 3099 ATGACCCACTACCGGCGATAAAGTGAGATTGGGCGATACCGACCTGATCTTTAGAAGTCG 3158
QY 991 AACATGACTATACCACTATCGGAGCACTTAATTTGGCGGGTAAAACTATCCGTG 1050
Db 3159 AACATGACTGCACCACTTATGCGGAAGAAATTAAGTTTGGTGGCGGTAAAAACCATTCGCG 3218

Db 187 AGCCACGAACCTCGATCTTGTGCTCACTAACGCCCTGATCGTGATTAACCCGGCATTTAT 246
Qy 1140 AAAGCCGACATTTGGGATTAATAACGCGCAAAATCCATGGCATTTGGCAAGGCGAGGAACAAG 1199
Db 247 AAAGCCGATTTGGCATTAATAATGCGAAATCCATGGCATTTGGCAAGGCGAGGAACAAG 306
Qy 1200 GACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAAGCACTAGCA 1259
Db 307 GACATGCAAGATGGCGTTTGCAACAATCTTTGGGTGGGCCCTGCTACTAGGGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATTAATACCGCTGGGGGAATCGATTCACACACCCACATTCCTTCTCCA 1319
Db 367 GCTGAAGGGGTGATTTGTCAGCTGGTGGGAATGACACCCACATTCACATTTATTTCTCC 426
Qy 1320 CAACAATTTCCCTTACCGCTCTAGCCAAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
Db 427 CAACAATTTCCACAGCATTTGCCAGCGGATCACAACCATGATTTGGTGGGGAAACAGGT 486
Qy 1380 CTTGTAGATGGCAGAAATGCGACTATCATCTCGGGCAAAATGGAACTTGGACCGCATG 1439
Db 487 CCAGCTGATGGGACTAACCGCATCTACCATCACTCCGGGGCGCTGGAACTTTAAACCCATG 546
Qy 1440 TTGGCGCAGCAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGGTCCCTCTGAAGAATATGCCATGAACCTTGGGCTATTTGGGTAAAGGAATGTGCT 606
Qy 1500 AGCAAAAAACAACTTGTAGAACAAATAGAACGGCGCGCATTTGTTTTAAATTCATGAA 1559
Db 607 TATGAACCTTCCCTGTGTCGATCACTCGAAGCTTGGAGCCATTTGGCTTTAAATCCAGAA 666
Qy 1560 GACTGGGGCAACAACAAGTGGGATCGATCACTGCTTGGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGGTAGCACACCTGCGAGCCATCTACCATTTGCTTGAATGTGGCTGACAAATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACAGGATACAGTCAATGAGGAGGTTATGTAGATGACACCTA 1679
Db 727 GTGCAAGTGGCTATCCACACCGATACCTTGAATGAAGCGGGCTGTGTGGGAAGACATTTG 786
Qy 1680 AATGCAATGAACGGCGCGCATCAATCGCTTACCACATTCAGGAGCGGGTGGAGGACAC 1739
Db 787 CAAGCCATTTGGTGGCGCATATCCACACTTTCACACTGAAGTGTGTGGGGCGGAC 846
Qy 1740 TCACCTGATGTTACCATGGCAGGCGAGCTCAATATTTCTACCTCTCCACACACCC 1799
Db 847 GCTCCGATGTCATTAAGATGCTGGCGAATTTAAACATCTCCAGCTTTACCAACCC 906
Qy 1800 ACTATTCCTATACCATTAATACGGTTGCGAACAACATTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACCATTCTTTTACCGTGAATACAGAAGCCGAAACACATGAGACATGTTGATGGTGGCCAC 966
Qy 1860 CACCTAGACAAACGATCCCGGAGGATTAACAATTTCTCAAGCCGTATCCGCCCGCGC 1919
Db 967 CACTTGGATAAAAACATCAAGAAGATGTCAGTTCCTGATTCCTAGGATTCGCCCCCAA 1026
Qy 1920 TCTATCCGGCTGAAGATGTCCTCCATGATATGGGTGTGATCGCGGATGACAAAGCTCGGAT 1979
Db 1027 ACCATCCGGCTGAGGACAACTCCAGATATGGGATTTCTCTATCACAGCTCTGAC 1086
Qy 1980 TCGAAGCAATTTGGCGTGCAGGCGAAGTGAATCTCTGAACTTGGCAGACTTGGCGATAAG 2039
Db 1087 TCCCAAGCGATGGCGGTGTAAGCGAGTCAACCCGCACTTGGCAAAACAGCGGACAAA 1146
Qy 2040 AATAAAAAGATTTGGTAAGCTTCTGMAAGTGGCAAGATACGATAATTTCCGCAAT 2099
Db 1147 AACAAAAAGATTTGGTTCGCTGAGGAAAAAGGCGATATGACAACTTCGCAATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTCACACGGCATTTCTGAATAT 1266
Qy 2160 ATCGGCTCTGGAAGGGGCAAGATCCCGCATTTGGTGGTGAATTCCTGCCCTTTT 2219
Db 1267 GTCGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTGTCTTTGGAGTCTCTGCTTTT 1326

RESULT 8

US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US2004241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PR5241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

Query Match 31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity 70.6%; Pred. No. 1.1e-195;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAACAAGAAATATGTAATACCTACGGACCCACCAAGGCGGATAAAGTCCG 959
Db 15 AAGATTTCCAGAAAGAATAATGTTCTATGATGTTCCCACTACCGGGATCGTGTAGA 74
Qy 960 TTAGGAGATACCGATCTTTGGGCAAGATAGAACTATACCACTATCGCGGAAGAA 1019
Db 75 CTCGCGACACTGATTTGATCTTAGAAGTGAGCATGATTTGCCACTTATGTTGAAGAG 134
Qy 1020 CTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTACAGCAATAGCCCTGAT 1079
Db 135 ATCAAAATTTGGGGCGGTAAAACTATCCGTGATGGGATGAGTCAAAACCAATAGCCCTAG 194
Qy 1080 GAAACACCTTAGATTTTAGTCATCACTAACGCGATGATTTATCGACTACACCGGATTTAC 1139
Db 195 TCTTATGAATAGATTTGGTGTCTACTAACGCCCTCATTTGTGACTATACGGGCAATTTAC 254
Qy 1140 AAAGCCGACATTTGGGATTTAAAAACGCGCAAAATTCATGTCATTTGGCAAGGACGAAACAAG 1199

[illegible]

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 251

; LENGTH: 1815

; TYPE: DNA

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (51) ... (1757)

US-09-895-913A-251

Query Match 30.7%; Score 885.8; DB 3; Length 1815;

Best Local Similarity 70.0%; Pred. No. 7.2e-192;

Matches 1193; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY	900	AAAAAGAAAAACAAGAAATATGTAATACCTACGGACCCACCAAGGCGATAAAGTGCGC	959
DB	57	AGATTAGCAGAAAGAATATGTTCTATGTATGGCCCTACTACAGCGGATTAAGTGAGA	116
QY	960	TTAGGAGATACCGATCTTTTGGGAGAGTAGAAACATGATATACCATCTATGGCGAAGAA	1019
DB	117	TTGGGCGATACAGACTTGATCGCTGAAGTAGAACAATGACTACACCATTTATGGCGAAGAG	176
QY	1020	CTTAAATTTGGCGGCTAAACTATCCGTGAGGATATGGGTACAGAGCAATAGCCCTGAT	1079
DB	177	CTTAAATTTGGCGGCTAAACCCCTGAGAGAGGATGAGGCAATCCCAACACCCCTAGC	236
QY	1080	GAACACACCTAGATTTAGTTCATCACTAACCGCATGATTTATGACTACACCGGATTTAC	1139
DB	237	AAAGAAGAATTGGAATCAATCACTAAGCTTTAATCGTGGATTACCGGATTTAT	296
QY	1140	AAAGCGCATTTGGGATTAATAACGGCAAAATTCATGGCATTTGGCAAGGAGGAAACAAG	1199
DB	297	AAAGCGGATTTGTTATTAAGATGSCAAAATCGCTGGCATTTGGTAAAGCGGTAACAAA	356
QY	1200	GACATGCAAGATGGCGTAAGCCCTCATATGCTGTGGGTGGGGACAGAGCACTAGCA	1259
DB	357	GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTAGGTCTGTACTTGAAGCCTTAGCC	416
QY	1260	GGGGAAGGTATCATTTATACCGCTGGGGAATCGATTTCACACACCACTTCCTTCTCCA	1319
DB	417	GGTGAAGGTTTATCGTAATCTGCTGGTGTATGACACACATCCATTCATTTCACCC	476
QY	1320	CAACAATTCCTTACCGCTTAGCCATGCGGTTTACCAACCATGTTTGGAGCGGCAACAGT	1379
DB	477	CAACAATTCCTTACAGCTTTTGAAGCGGTGTAAACCAACCATGATTGGTGGCGGAATG	536
QY	1380	CTGTAGATGGCAAGATCGACTACTATCATCTCGGGCAATGGAATTTGCACCGCATG	1439
DB	537	CCTGCTGATGGCACTAAATGCGACTACTATCACTCCAGGCGAAGAAATTTAAAAATGGATG	596
QY	1440	TTGCGGCGACAGAGAGTATTTCTATGAATGTGGCTTTTGGGCAAGGCAATAGCTCT	1499
DB	597	CTCAGAGCGGCTGAAGAATATTTCTATGAATCTAGGTTTCTGGCTAAAGGTAAACGCTCT	656
QY	1500	AGCAAAAAACAATTTGTAGAACCAAGTAGAAGCGGCGGATTTGTTTAAATTTGCATGAA	1559
DB	657	AACGAGCGGAGCTTAGCCGATCAATTTGAAGCTGGTGGATTTGCTTTAAATCCAGAA	716
QY	1560	GACTGGGCGACAAACCAAGTCGATCGATCTGCTTGAGCGGTGGCAGATGAATACGAT	1619
DB	717	GACTGGGCGACCACTCTTCTGCAATCAATCATCGTTAGATTTTCAGACAAATACGAT	776
QY	1620	GTGCAAGTTTGTATCCACACCATACAGTCAATGAGCGAGGTATGTAGATGACACCCTA	1679
DB	777	GTGCAAGTCGCTATCCACACAGACATTTTGAATGAAGCGGTTTGGGTGGAAGACACTATG	836
QY	1680	AATGCAATGAACGGGCGGCCATCCATGCTACCATATTGAGGAGCGGTTGGAGGACAC	1739
DB	837	GCAGCTATTGGCGGACGCACTATGACACTTTCCACACTGAAGGTGCTGGCGCGGACAC	896
QY	1740	TCACCTGATTTATCACCATGGCAGCGAGCTCAATTTCTACCCCTCTCCACCAACCC	1799
DB	897	GCTCCTGATTTATTAAGTAGCTGGTGAACACAACTTTCTTCCCGCTTCCACTAACCC	956

QY	1800	ACTATTCCTTATACCAATTAATACGTTTGCAGAACACTTAGACATGCTCATGACATGCCAC	1859
DB	957	ACTATTCCTTATCACTGTGAATACAGAAGCAGAACACATGGACATGCTTATGGTGGCCAC	1016
QY	1860	CACCTAGACAAACGCATCCCGGAGGATTTCAATTTTCTCAAGCCGATTCGGCCCGGC	1919
DB	1017	CACCTGGATAAAGCATTAAGAAGATGTTTCAGTTCGCTGATTCAGGATTCGGCCCTCAA	1076
QY	1920	TCATTCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGAT	1979
DB	1077	ACATTTGGGCTGAGACACTTTGCGATGACATGGGGATTTTCTCAATCACCAGCTCTGAC	1136
QY	1980	TCGCAAGCAATGGGCGGTGAGGCAAGTATTCTCTCGAACTTGGCAGACTGCGGATAG	2039
DB	1137	TCTCAAGCTATGGTGTGGTGAAGTTTATCACTAGAACTTGGCAACAGCTGACAAA	1196
QY	2040	AATAAAGAAATTTGGTAAGCTTCTGGAAGATGGCAAGATAAGATAATTTCCGCTT	2099
DB	1197	AACAAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGGCGATACGACACTTCAGGATC	1256
QY	2100	AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCCGGCTGAGCGAGTAT	2159
DB	1257	AAAGCTACTTGTCTAAATACCAATTAACCCAGGATCGCTCATGGATTTAGCGAGTAT	1316
QY	2160	ATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGGTGGTGTGGAATTCCTGCTTTT	2219
DB	1317	GTAGGTTCTGTAGAGTGGGCAAGTGGCTGACTTGGTATTGTGGAGTCCCGCATTTT	1376
QY	2220	GGCGTAAACCCAAATTCGTGATCAAGGCGGTATGGTGTCTTCTGTGAATGGGCGAT	2279
DB	1377	GGCGTAAACCCCAACATGATCATCAAGGCGGTTCTATGCGTTGAGTCAAAATGGGTGAC	1436
QY	2280	TCATAACGCTCTGTGCCCTCCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCAAC	2339
DB	1437	GCAACGCTTCTATTCCTCCCAACCAAGTATTATACAGAGAAATGTTGCTCATCAT	1496
QY	2340	GGCAAGCGGAAATTTGACACAGCATCACTTTTGTTCCTCAAGTCCGCTATGAAAATGGC	2399
DB	1497	GGTAAGCCAAATACGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGC	1556
QY	2400	GTCAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGTCAAAAATCGCGTAAACATC	2459
DB	1557	ATTAAGAGAAATTTAGGCTTGAAGACAAAGTGTGCGGTAAAAAATTCAGAAACATC	1616
QY	2460	ACCAAGAAAGCTTCAAGTTCAAGCAAAAACGCAAAATCAACGTCGATCCGAAACCC	2519
DB	1617	ACTAAAAAGACATGCAATTTCAACGACACTACCGCTCACATTTGAAGTCAATCCTGAAACT	1676
QY	2520	TTGAGGCTCTTGTAGATGCAAACTCTGCACTCTTAAACCCACTCGCAAGTGCCTCTA	2579
DB	1677	TACATGTGTTCTGGATGCAAGAAAGTAACTTCTTAACCGCAATAAAGTGAAGCTTG	1736
QY	2580	GCCAGCGCTACACTTTTCTTAGG	2604
DB	1737	GCGCAACTCTTTAGCATTTTCTAGG	1761

RESULT 11

US-10-282-122A-22427

; Sequence 22427, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

```

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22427
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-282-122A-22427

```

Query Match	30.7%	Score	884.8	DB	7	Length	1710
Best Local Similarity	70.0%	Pred. No.	1.2e-191				
Matches	1192	Conservative	0	Mismatches	512	Indels	0
Gaps	0						
Qy	900	AAATGAAAAACNAGAAATATGTAATACCTACGGACCCCAACCAAGCGGATAAAGTCGC	959				
Db	7	ARAGATTAGCAGAAAGAATATGTTCTATGTATGGCCCTACTACAGCGGATAAAGTGAGA	66				
Qy	960	TTAGGAGATACCGATCTTTGGCGAGAAGTAGAACATGACTATATACCCTATGCGGAAGAA	1019				
Db	67	TTGGCGGATACAGACTTGATCGCTGAAGTAGAACATGACTACACCATTTATGGCGAAGAG	126				
Qy	1020	CTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGACCAATAGCCCTGAT	1079				
Db	127	CTTAAATTCGGTGGCGGTAAAAACCTTGAGAGAGGCGATGAGCCAAATCCAACACCCCTAGC	186				
Qy	1080	GAAAAACACCTAGATTTAGTCATCACTAAACCGGAGATGATTCGACTACACCGGATTTAC	1139				
Db	187	AAAGAAGAAATTGGATCTAAATCATCACTAAACGCTTTAATCGTGATTACACCGGTATTTAT	246				
Qy	1140	AAAGCCGACATTTGGGATTTAAAAACGGCAAAATTCATGGCATTTGCAAGGCGAGAAACAAG	1199				
Db	247	AAAGCGGATATTGGTATTTAAAGATGGCAAAATCGCTGGCATTTGTTAAGCGGTAAACAA	306				
Qy	1200	GACATGCAAGATGGCGTTAAGCCCTCATATGGTCTGTGGTGTGGGCACAGAGCACATAGCA	1259				
Db	307	GACATGCAAGATGGCGTTAAAAACAATCTTTAGCGTAGGTCTCTGCTACTGAAGCCTTAGCC	366				
Qy	1260	GGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACACACCACTTCTCTTCTCCCA	1319				
Db	367	GGTGAAGGTTTGATTCGTAACATGCTGTGGTGATTTGACACACACATCCACTTCATTTACACC	426				
Qy	1320	CAACAAATTCCTACCGCTCTAGCCAAATGGCGGTTAACACCATGTTTGGAGCGGCACAGGT	1379				
Db	427	CNACAAATCCCTACAGCTTTTGGCAAGCGGTGAACACCATGATTTGTTGGCGGAACTGGT	486				
Qy	1380	CCTGTAGATGGCAAGAAATGGCACTATCATCTCCGGGCAAAATGGAACTTTGCACCCGCAATG	1439				

487	CTGTGATGGCACTTAATCGGACTACTATCACTCCAGGCAGAAGAAATTTAAATGGATG 546	Db
1440	TTCCGCGCAGCAGAGAGATTTCTATGAATGTGGCTTTTTGGGCAAGCAATAGCTCT 1499	Qy
547	CTCAGAGCGCTGAAGAAATATTCTATGAACCTTAGTGTTCCTGGCTAAAGTAAAGCTTCT 606	Db
1500	AGCAAAAAACAACCTTGTAGNACAAGTAGAAGCGGGCGCATTTGGTTTTAAATTCGATGA 1559	Qy
607	AACGACGCGAGCTTAGCCGATCAAAATGGAAGCTGGTGCATTTGGCTTTAAAAATCCAAG 666	Db
1560	GACTGGGGCACAACAACCAAGTGCATCGATCACTGCTTGAGCGGTGCAGATGAATACGAT 1619	Qy
667	GACTGGGGCACCACCTCTTTCGAATCAATCATGCGTTAGATGTTGCAGACAATACGAT 726	Db
1620	GTCCAAATTTGTATCCACACCGCATACAGTCAATAGAGCGAGTTATGTAGATGACACCTTA 1679	Qy
727	GTGCAAGTCGTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGAAGACACTATG 786	Db
1680	AATGCAATGAACGGGCGGCCATCCATGCTTACCACATTTGAGGGAGCGGTGGAGACAC 1739	Qy
787	GCAGCTATTTCGCGGACGCACTATGSCACATTTCCACACTGAAGGTGCTGGCGCGGACAC 846	Db
1740	TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACCCCC 1799	Qy
847	GCTCTCGATATTATTAAGTAGCTGTGTGAACACACATTTCTCCGCTTCCACTAACCCC 906	Db
1800	ACTATTCCCTATACCAATTAAATACGGTTGCAGAACACTTAGACATGTCTCATGACATGCCAC 1859	Qy
907	ACTATCCCTTTCACCTGTGAATACAGAAAGCAGAACACATGACATGCTATTATGTTGCCAC 966	Db
1860	CACCTAGACAAACGGCATCCGCGAGGATTTACAATTTTCTCAAGCCGTATCCGCCCGGC 1919	Qy
967	CACCTTGGATAAAAGCAATTAAAGAAGATGTTCAGTTCCGCTGATTCAAAGGATTCGGCCCTCAA 1026	Db
1920	TCTATCGCGCTCAAGATGTGCTCAATGATATGGGTGTGATCCGGATGACAAGCTCGGAT 1979	Qy
1027	ACATTTGCGCTGAAACACACTTTGCAATGACATGGGATTTTCTCAATCACCGACTGTGAC 1086	Db
1980	TCGCAAGCAATGGGCGGTGCAGCGAAGTGATTCCTCGAACTTTGGCAGACTGCGGATAAG 2039	Qy
1087	TCTCAAGCTATGGTCTGTGGGTGAAGTTATCACTAGAACCTTTGGCAACAGCTGACAAA 1146	Db
2040	AATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGTGGCAAGATAACGATAAATTTCCGAT 2099	Qy
1147	AACAAAAAAGAAATTTGGCCGCTTTGAAAGAGAAAAAGGCGATTAACGACAACTTCAGGATC 1206	Db
2100	AAGCGTACATCTCCAAATACATCTCAACCCCGCTTTGACCCACGCGGTGAGCGAGTAT 2159	Qy
1207	AAACGCTACTTGTCTAAATACACATTAACCCAGCATCGCTCATGGATTAGCGAGTAT 1266	Db
2160	ATCGGCTCTGTGAAGAGGGCAAGATCGCCGACTTGGTGGTGTGGAATCCTGCTTTTTT 2219	Qy
1267	GTAGGTTCTGTAGAAGTGGCAAGTGGCTGACTTGGTATTGTGGAGTCCCGCATTTT 1326	Db
2220	GGCGTAAACCCCAAAATFCGTGATCAAGGCGGTATGGTGGTCTTCTCTGAAATGGGGCAT 2279	Qy
1327	GGCGTAAACCCCAACATGATCAACAAAGCGGGTTCATTGCGTTGAGTCAAAATGGGTGAC 1386	Db
2280	TCTAAACGCTCTGTGCCACTCCCAACCGTTTTATTACCGCAAAATGTTTGGGCATCAC 2339	Qy
1387	GCGAAGCGTTCTATCCCTACCCCAACCAAGTTTATTACAGAGAAATGTTTCGCTCATCAT 1446	Db
2340	GGCAAGCGGAAATTTGACACACGATCACATTTTTTTTTCCAAAGTCGCGCTATGAAATGGC 2399	Qy
1447	GGTAAGCCAAATACGATGCAACATCACTTTTGTCTCAAGGGCTTATGACAAAGGC 1506	Db
2400	GTGAAGAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAACTGCCGTAAACATC 2459	Qy
1507	ATTAAGAAGAAATTAGGGCTTGAAGAACAAGTGTTCGCGGTAAAAAATTCAGAAAAATC 1566	Db
2460	ACCAAGAAAGACTTCAAGTTTCAACGACAAACGGCAAAATCAACCGTCGATCCGNAACC 2519	Qy
1567	ACTAAAAAAGACATGCAATTTCAACGACACTACCGCTCACATGGAAGTCAATCTCGAACT 1626	Db

QY 2520 TTGAGGCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTSCCTCTA 2579
Db |||||
1627 TACCATGTGTCTGTGATGGCAAGAACTTCTAAACCGCAATAAAGTGAGCTTG 1686
QY 2580 GCCCAGCGCTACACTTTCTTCTAG 2603
Db |||||
1687 GCGCAACTCTTTAGCATTTTCTAG 1710

RESULT 12

US-10-500-447A-5
; Sequence 5, Application US/10500447A
; Publication No. US20050150016A1
; GENERAL INFORMATION:
; APPLICANT: PARK, Hee-Sung
; TITLE OF INVENTION: Method for producing a recombinant protein using pollen
; FILE REFERENCE: YLOP040518US/PCT
; CURRENT APPLICATION NUMBER: US/10/500,447A
; CURRENT FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: KR 2001-71712
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-500-447A-5

Query Match 30.5%; Score 880; DB 9; Length 1710;
Best Local Similarity 69.8%; Pred. No. 1.5e-190;
Matches 1189; Conservative 0; Mismatches 515; Indels 0; Gaps 0;

QY 900 AAAAAAGAAACAGAAATATGTAATACCTAGCGAACCCACCAAGCGGATTAAGTGGCG 959
Db |||||
7 AAGATTAGCAGAAAGAAATATGTTTCTATGTATGGCCCTACTACAGCGGATTAAGTGAGA 66
QY 960 TTAGGAGATACCGATCTTTGGGAGAGATAGAACATGACTATACCACTATGGCGAAGAA 1019
Db |||||
67 TTGGGCGATACAGACTTGTATCGCTGAAGTAGAACATGACTACACCAATTTATGGCGAAGAG 126
QY 1020 CTTAAATTTGGCGCGGTAAACCTATFCGTGAGGGTATGGGTACAGAGCAATAGCCCTGAT 1079
Db |||||
127 CTTAAATTTGGCGCGGTAAACCTTAAAGAGGATGAGCCATTAACAACTAGC 186
QY 1080 GAAACACCTTAGATTAATGATCACTAACCGGATGATTAATGATCACTACCGGGATTTAC 1139
Db |||||
187 AAAGAAGAACTGGATCTAATCACTAAGCGCTTTAATCGTGGATTAACACCGGTATTTAT 246
QY 1140 AAAGCCGACATTTGGATTAAGAAACCGCAAAATCCATGGCATTTGGCAAGCGCAAGAAACAG 1199
Db |||||
247 AAAGCGGATTTGGTATTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGTAACAA 306
QY 1200 GACATGCAAGATGGCGTAAAGCCCTCATATGTTGTTGGGTGTTGGGACAGAGCACTAGCA 1259
Db |||||
307 GACATGCAAGATGGCGTTAAACAACTCTTAGGTGGTCTGCTACTGAAGCTTAGCC 366
QY 1260 GGGGAAGGTATGATTAATACCGTGGGGAATGATTTCAACACCACTTCCTTTCTCCA 1319
Db |||||
367 GGTGAAGTTTGTATCGTAATCTGCTGGTGTATTGACACACACATCCACTTCATCTCCCCC 426
QY 1320 CAACAAATTCCTACCGCTTAGCCAATGGGTTTACAACTGTTTGGAGCGCGCACAGGT 1379
Db |||||
427 CAACAAATTCCTACAGCTTTTGAAGCGGTAAACCAACGATGTTTGGTGGCGAACTGGC 486
QY 1380 CTTGTAGATGGCAGCAATCGGACTACTATCACTCCGGGAAATGGAATTTGCACCGCATG 1439
Db |||||
487 CTTGTGTAGTGGCACTAACCAACCACTATCACTCCAGGTAGAGAAATTTAAATGGATG 546
QY 1440 TTGCGCGCAGCAGAGATTAATCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db |||||
547 CTCAGAGCGGCAGAGAAATTTCTATGAATTTAAGTTTCTTAGCTAAAGGTAAACGCTTCT 606

QY 1500 AGCAAAAAACAACCTGTAGAACAAAGTAGAACGGGCGCGATTTGGTTTAAATTCGATGAA 1559
Db |||||
607 AACGATGCAAGCTTAGCCGATCAAAATTGAAGCGGTCGATTTGGCTTTAAAAATCCACGAA 666
QY 1560 GACTGGGGCAACAACCAAGTCGATCACTGCTTGGAGCGTGGCAGATGAATACGAT 1619
Db |||||
667 GACTGGGGCACCACTCTCTTCTGCAATCAATCATGCGTTAGATGTTGCGGACAAATACGAT 726
QY 1620 GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTA 1679
Db |||||
727 GTGCAAGTCGCTATCCACACAGACATTTGAAAGAGCCGGTTGTGTAGAAGACACTATG 786
QY 1680 AATGCAATGAACGGGCGCCATCCATGCCCTACACATTTGAGGGAGCGGGTGGAGGACAC 1739
Db |||||
787 GCAGCCATTCGCGGACGCACTATGCACTTCCACACTGAAGCGCTGGTGGCGGACAC 846
QY 1740 TCACCTGATGTTATACCATGGCAGGCGAGCTCAATATTCTACCTCTCCACACCCCTC 1799
Db |||||
847 GCTCTGATATTATTAAAGTAGCTGGTGAACACAACTTCTGCCCGCTTCCACTAAACCCC 906
QY 1800 ACTATTCCCTATACCATTAATACGGTTGAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db |||||
907 ACTATCCCTTCACTGTGAATACAGAGAGCAACATGGACATGCTTATGGTGTGCCAC 966
QY 1860 CACTAGACAAACGCACTCCGCGAGGATTTAACAATTTTCTCAAGCCGCTATCCGCCCCGCGC 1919
Db |||||
967 CACTTGGATAAAGCATTAAGAAGATGTTCACTTCGCTGATTCAGGATCCGCCCTCAA 1026
QY 1920 TCTATCGCGCTGAAGATGTCTCATGATATGGTGTGATCGCGATGACAAAGCTCGGAT 1979
Db |||||
1027 ACTATTGCGGCTGAGACACTTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC 1086
QY 1980 TCGCAGCAATGGGGCTGCAGCGGAGTATTCCTCGAACTTGGCAGAGCTGCCGATAG 2039
Db |||||
1087 TCTCAAGCTATGGGTCGTGGGTGAAGTTATCACAGAACTTGGCAACAGCTGACAAA 1146
QY 2040 AATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGGCAAGATTAACGATAATTTCCGCAAT 2099
Db |||||
1147 AACAAAAAGAAATTTGGCCGCTTGAAGAGAAAGGCGATTAACGACACTTCAGGATC 1206
QY 2100 AAGCGCTACATCTCCAAATACATAACAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
Db |||||
1207 AAACGCTACTTGTCTAAATACACCAATTAACCCAGCGATCGTCAATGGGATTTAGCGAGTAT 1266
QY 2160 ATCGGCTCTGTGAAAGGAGCAAGATCGCCGACTTGGTGGTGGATCTCTGCCCTTTTT 2219
Db |||||
1267 GTAGGTTCTGTAGAAGTGGCAAAAGTGGCTGACTTGGTGTGGAGTCCCGCAATCTTT 1326
QY 2220 GGCGTAAACCCCAAAATCGTGATCAAGGGCGGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
Db |||||
1327 GGCGTGAACCCCAAAATGATCATCAAGGGCGGATTCATTCGATGATCAATGGGTGAT 1386
QY 2280 TCTAACCGCTGTGTGCCCATCTCCCAACCGGTTTATTAACCGCGAAATTTTGGGCATCAC 2339
Db |||||
1387 GCGAACGCTTCTATCCCTACCCCAACCGGTTTATTAATAGAGAAATGTTTCGCTCATCAT 1446
QY 2340 GGCAAGCGGAAATTTGACACCGACATCACTTTTGTTCGAAAGTCCGCTATGAAAATGGC 2399
Db |||||
1447 GGTAAAGCTAAATPACGATGCAACATCACTTTTGTGCTCAAGCGGCTTATGACAAAGGC 1506
QY 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGCGTAAACATC 2459
Db |||||
1507 ATTAAGAAAGAAATTAGGGCTTGAAGCGAAGTGTGGCGGTAAAAAATTTGAGAAACATC 1566
QY 2460 ACCAAGAAAGACTTCAAGTTTCAACGCAAAACCGCAAAAATCAACCGTCGATCCGAAACCC 2519
Db |||||
1567 ACTAAAAAGACATGCAATTAACGACACTACCGCTCACAATTGAAGTCAATCTCTGAAACT 1626
QY 2520 TTGAGAGTCTTTGTAGATGCGAAACTCTGCACTTAAACCCACCTCGCAAGTGCCTCTA 2579
Db |||||
1627 TACCATGTGTCTGTGATGGCAAAAGAACTTCTTAACCCAGCGCAATAAAGTGAGCTTG 1686

Qy	2580	GCCAGCGCTACACTTCTTCTCTAG	2603
Db	1687	GCACAACTCTTTAGCAATTTCTAG	1710
RESULT 13			
US-10-476-313-11			
; Sequence 11, Application US/10476313			
; Publication No. US20040241175A1			
; GENERAL INFORMATION:			
; APPLICANT: MCKENZIE, BRENT			
; APPLICANT: BOYLE, JEFFEREY			
; APPLICANT: LEW, ANDREW			
; TITLE OF INVENTION: Antigen Targeting			
; FILE REFERENCE: BDWP-002			
; CURRENT APPLICATION NUMBER: US/10/476,313			
; CURRENT FILING DATE: 2003-11-0-24			
; PRIOR APPLICATION NUMBER: PR5241			
; PRIOR FILING DATE: 2001-05-25			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: Patentln version 3.1			
; SEQ ID NO 11			
; LENGTH: 1717			
; TYPE: DNA			
; ORGANISM: Helicobacter pylori			
US-10-476-313-11			
Query Match 30.4%; Score 876.8; DB 8; Length 1717;			
Best Local Similarity 69.7%; Pred. No. 8.1e-190;			
Matches 1187; Conservative 0; Mismatches 517; Indels 0; Gaps 0;			
Qy	900	AAAATGAAAAAACAAGAATATGTAAATACCTACGACCCCAACAAAGCGGATAAAGTGC	959
Db	13	AAGATTAGCAGAAAGAATATGTTCTATGTATGGCCCTACTACAGGTGATAAAGTGAGA	72
Qy	960	TTAGGAGATACCGATCTTTGGGAGAAGTAGAATGACATCATATACCACCTATGGCAAGAA	1019
Db	73	TTGGCGGATACAGACTTGTGCGTGAAGTAGAAATGATGACTACACCAATTTATGGCGAAG	132
Qy	1020	CTTAAATTTGCGCGGGTAAACATCATCGTGAAGGTATGGTTCAGAGCAATAGCCCTGAT	1079
Db	133	CTTAAATTCGTGGCGGTAAACCTTAAGAGAAGCATGAGCCNAATCTAACACCTTAGC	192
Qy	1080	GAAACACCCCTAGATTTAGTTCATCACTAAACCGCATGATTAATCGACTACACCGGATTTAC	1139
Db	193	AAAGAAGAACTGGATCTAATCATCACTAAACGTTTAATCGTGATTACACCGGTATTTAT	252
Qy	1140	AAAGCCGACATTTGGGATTTAAAAACGGCAAAATTCATCGGCATTTGGCAAGGCGAGAAACAAG	1199
Db	253	AAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGGCGGTAAACAA	312
Qy	1200	GACATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGTGGGCAAGAGCACTAGCA	1259
Db	313	GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTGGGTCTGTCTAATGAAGCCCTTAGCC	372
Qy	1260	GGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTACACACCCACTTCTTCTTCCCA	1319
Db	373	GGTGAAGTTTGATCGTAACTGCTGGTGGTATTGACACACATCCACTTCATTTTCAACC	432
Qy	1320	CAACAAATTCCTTACCGCTCTAGCCAATGGCGGTACAAACCAATGTTTGGAGGCGGACAGGT	1379
Db	433	CAACAAATTCCTACAGCTTTTGGCAAGCGGTGTAAACAACCATGATTTGGTGGCGGAATCTGGT	492
Qy	1380	CCTGTAGTGGCAGGATGCACTACTACTCTCGGGCAAAATGGAATCTTGACCCGATG	1439
Db	493	CCTGCTGATGGCACTAAACGCGACTACTATCACTCCAGGTGAAGAAATTTAAATATGATG	552
Qy	1440	TTGGCGCGCAGCAGAAAGTATTTCTATGAATGCGGCTTTTTTGGGCAAGGCAATAGCTCT	1499
Db	553	CTCAGAGCGGCTGAAGATATTTCTATGAATTTAGGTTTCTTGGCTAAAGGTAAAGCTTCT	612
Qy	1500	AGCAAAAAACAATTTGTAGAAACAATGAGACGGCGCGATTTGGTTTTAAATTCATGAA	1559

Db	613	AACGATGCGAGCTTACCGCATCAAAATTGAAGCCGGTGCATTTGGCTTTTAAAAATTACAGAA	672
Qy	1560	GACTGGGGACAAACCAAGATGCGATCGATCACTGCTTGAGCGTGCAGATGAATACGAT	1619
Db	673	GACTGGGGACCACTCTTCTGCAATCAATCATGCGTTAGATGTTGGGACAAATACGAT	732
Qy	1620	GTGCAAGTTTGTATCCACACCCGATACAGTCAATGAGGCGAGGTATGTAGATGACACCCCTA	1679
Db	733	GTGCAAGTCGTATCCACACAGACACCTTTGAATGAAGCCGGTTGTGTAGAAGACACTATG	792
Qy	1680	AATGCAATGAACGGCGCGCATCCATGCTCCACCAATGAGGAGCGGTGAGGACAC	1739
Db	793	GCTGCTATTGCTGGAGCGCATATGCACATTTTCCACATGGAAGCGCTGGCGCGGACAC	852
Qy	1740	TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACCCCC	1799
Db	853	GCTCCTGATATTATTAAGTGAAGCGGTGAACACAACTTCTCCCGCTTCCACTAACCCCC	912
Qy	1800	ACTATTCCCTATACCAATTAATACGGTTGCGAAGACATTAGACATGCTCATGACATGCCAC	1859
Db	913	ACCATCCCTTTCACCGTGAATACAGAAAGCAGACCATGAGACATGCTTATGTTGTGCCAC	972
Qy	1860	CACCTAGACAAAACGCAATCCGCGAGGATTTACAATTTTCTCAAAGCGTATCCGCCCCGGC	1919
Db	973	CATTGGATTAAGCAATTAAGAAAGATGTTTCAGTTTCGTGATTAAGGATCCGCCCTCAA	1032
Qy	1920	TCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGAT	1979
Db	1033	ACCATTGCGGCTGAAGACACTTTTGATGATCATGGGATTTTCTCAATCACCAGTCTGTAC	1092
Qy	1980	TGCGAAGCAATGGGGGTGACGGCGAAGTGATTCCTCGAACTTGGCAGAGCTGCGGATTAAG	2039
Db	1093	TCTCAAGCGATGGCGCTGTGGGTGAAGTTATCACTAGAACTTTGGCAAAACAGCTGACAA	1152
Qy	2040	AATTAATAAGAAATTTGCTGAAGCTTCTGGAAGTGGCAAGATAACGATAATTTCCGCATT	2099
Db	1153	AACAAGAAAGAAATTTGGCCGCTTGAAGAAGAAAAAGCGGATTAACGAACTTCAGATC	1212
Qy	2100	AAGCGTTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGACGCGAGTAT	2159
Db	1213	AAAGCTACTTGTCTAAATACCAATTAACCCAGCGATCGTCTATGGGACTACGGAGTAT	1272
Qy	2160	ATCGGCTCTGTGAAGAGGCAAGATCGCGCACTTGGTGGTGTGGAATCTCTGCTCTTTT	2219
Db	1273	GTGCGTCTCTAGAAGTGGGCAAGTAGCTGACTTGGTATTGTGGAGTCCAGCATTTCTTT	1332
Qy	2220	GGCGTAAACCCAAATCGTGATCAAGCGGTATGGTGTGCTTCTCTGAAATGGCGGAT	2279
Db	1333	GGCGTGAACCTTAACATGATCATAAGGTGGGTTCATTGCAATTAAGCCAAATGGCGCAT	1392
Qy	2280	TCTAACCGCTCTGTGCCACTCCCCAAACCGGTTTATTACCGCAATGTTTGGGCACTCAC	2339
Db	1393	GCGNACGTTCTATCCCTACCCCTCAACCGGTTTATACAGAGAAATGTTGCTCTCATCGT	1452
Qy	2340	GGCAAGCGAAATTTGACACACGATCACTTTTGTTCCTGAAAGTGCCTATGAAATGGC	2399
Db	1453	GGTAAAGCTAAATACGATGCAACATCACTTTTGTGCTCAAGCGGCTTATGCAAAAGGC	1512
Qy	2400	GTGAAGAAAGAGCTGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATC	2459
Db	1513	ATTAAAGAGAAATTAGGACTTGAAGACAAAGTGTGCCGGTTAAAAATTCGAGAAACATC	1572
Qy	2460	ACCAAGAAAGACTTCAAGTTTCAACGACAAACCGCAAAAATCACCGTTCGATCCGAAAAACC	2519
Db	1573	ACMAAAAGACATGCAATTAACGACACTACCGCTCAGATTGAGTCAATCTCTGAAACT	1632
Qy	2520	TTGAGGTCTTTGTAGATGGCAAACTCTGACCTCTAAACCCACCTCGCAAGTGCCTCTA	2579
Db	1633	TACCATGTGTTGTTGGATGCGAAAGAGTAACCTTCTAAACCGAGCTAATAAAGTGAGCTTG	1692
Qy	2580	GCCAGCGCTACACTTTCTTCTAG	2603
Db	1693	GCGCACTCTTTAGCATTTTCTAG	1716

RESULT 14

US-09-402-100-1
; Sequence 1, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter pylori
; FILE REFERENCE: 0136/06140
; CURRENT APPLICATION NUMBER: US/09/402,100
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1..)
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

Query Match 29.7%; Score 855.6; DB 3; Length 2385;
Best Local Similarity 69.8%; Pred. No. 6.8e-185;
Matches 1169; Conservative 0; Mismatches 504; Indels 1; Gaps 1;
QY 900 AAAATGAAACAAAGAAATATGTAATACCTACGGACCCACAAAGGCGATAAAGTGCGC 959
DB 7 AGATTAGCAGAAAGAAATATGCTTCTATGATGGCCCTACTACAGCGCATAAAGTGAGA 66
QY 960 TTAGGAGATACCGATCTTTGGCAGAGTAGAACATGACTATACCACTTATGGCGAAGAA 1019
DB 67 TTGGCGATACAGACTTGTATCGCTGAAGTAGAACATGACTACACCATTTATGCTGAAGAG 126
QY 1020 CTTAAATTTGGCGCGGTAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
DB 127 CTTAAATTTGGCGCGGTAAACCCCTTAAGAGAGGCAATGAGCCAAATCTAACCAACCTAGC 186
QY 1080 GAAACACCCCTAGATTAGTATCATCTAACGCGATGATTATGACTACACCGGGATTAC 1139
DB 187 AAAGAAGAACTGGATCTAATCATCTAACGCTTTAATCGTGATTACACCGGTAFTTAT 246
QY 1140 AAAGCCGACATTTGGGATTAAAAACCGCAAAATCCATGGCAATTGGCAAGCGCAGAAACAAG 1199
DB 247 AAAGCGATTTGTTATTTAAAGTAGGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAA 306
QY 1200 GACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGTGGGTGGGACAGAAAGCACTAGCA 1259
DB 307 GACACGCAAGATGGCGTTAAAAACAATCTTAGCGTGGGTCTGCTACTGAAGCCTTAGCC 366
QY 1260 GGGGAGGTATGATTATTAACCGCTGGGGAATCGATTACACACCCACTTCTTTCTCCA 1319
DB 367 GGTGAAGGTTTGAATGTAATCTGCTGGTGTATTTGACACACATCCACTTCATCTCCCCC 426
QY 1320 CAACAAATCCCTACAGCTTTTGAAGCGGTGTAAACACCATGATTGGTGGCGGCACTGAC 486
DB 427 CAACAAATCCCTACAGCTTTTGAAGCGGTGTAAACACCATGATTGGTGGCGGCACTGAC 486
QY 1380 CTTGTAGATGGCAGAAATGCGACTATCATCTCCGGGCAAAATGGAACCTTGCACCGCATG 1439
DB 487 CTTGTGTATGGCACTAAACGCAACCACTATCACTCCAGGTAGAAGAAATTTAAAAATTTCA 546

QY 1440 TTGGCGCAGCAGAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
DB 547 CTCAGAGCGGCTGAAGAAATATTCTATGAATTTTGGTTTCTTGGCTAAAGGTAACGCTTCT 606
QY 1500 AGCAAAAAACAACTTGTAGAACAGTAGAAGCGCGCGATTGGTTTAAATTTGATGAA 1559
DB 607 AACGATGCAAGCTTAGCCGATCAAAATTTGAAGCTGTGCGATTGGCTTTAAATTCACGAA 666
QY 1560 GACTGGGGCAACAACCAAGTGGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
DB 667 GACTGGGGCAACCACTCTTCTGCAATCAATCATGCGTTAGATTTTGGGACAAATACGAT 726
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGACGATTATGATAGTACACCCCTA 1679
DB 727 GTGCAAGTGTCTATCCACACAGACACTTTGAATGAAGCCGCTTGGTGGGAGACACTATG 786
QY 1680 AATGCAATGAACCGGGCGCCATCCATGCTTACCACTTGAAGGAGCGGGTGGAGGACAC 1739
DB 787 GCAGCTAATTGCGGACGCACTATGCACACTTACCACTGAAGCGCTGGCGGGGACAC 846
QY 1740 TCACCTGATGTTATCACCATGGCGGAGCTCAATATTCTACCTCTCTCCACACCCCTC 1799
DB 847 GCTCTGATATTATTAAGTGGCGGTGAACACACACATCTTACCCTTCCACTAACCCCT 906
QY 1800 ACTATTCCCTATACCAATTAATACGGTTTGCAGAACACTTAGACATGCTCATGACATGCCAC 1859
DB 907 ACTATCCCTTTACCGCTGAATACAGAAAGCGCAACACATGGACATGCTTATGGTGTGCCAC 966
QY 1860 CACTAGACAAACGATCCCGGAGGATTTACAAATTTTCTCAAGCCGCTATCCGCCCGGC 1919
DB 967 CACTTGGATAAAAGCATTAAGAAGATGTCAGTTGCTGATTCAAGGATTCGCCCTCAA 1026
QY 1920 TCTATCGCGCTGAAGATGCTCCATGATATGGTGTGATCGGATGACAAAGCTCGGAT 1979
DB 1027 ACCATTGCGGCTGAAGACACTTTTGCATGACATGGGATTTTCTCAATCCTAGTCTTGAC 1086
QY 1980 TCCAGACCAATGGGGCTGCAGCGCAAGTATCTCCTGAACTTGGCAGACTGCGGATGAAG 2039
DB 1087 TCTCAGCGATGGCGCTGTGGGTGAAGTTATCACTAGAACTTGGCAAAACAGCTGACAAA 1146
QY 2040 AATAAAAAAGAAATTTGGTAAAGTCTTCTGAAAGTGGCAAAAGATAAACGATTAATTTCCGCAAT 2099
DB 1147 AATAAAAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGGCGATACGACAACTTCAGGATC 1206
QY 2100 AAGCGTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGGGTGAGCGAGTAT 2159
DB 1207 AAACGCTACTTGTCTAAATACACCAATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT 1266
QY 2160 ATCGGCTCTGTGGAAGGGAAGATCGCCGACTTTGGTGTGTGGAATCTTGCCTTTTTT 2219
DB 1267 GTCGGTCTGTAGAGTGGGCAAGTGGCTGACTTGGTATTTGGAGTCCCGCATTTCTT 1326
QY 2220 GCGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
DB 1327 GGTGTGAAACCCAAATGATCATCAAGGCGGTTCATCGCATTTAGTCAAAATGGGTGAT 1386
QY 2280 TCTAAAGCGCTGTGGCCACTCCCAACCGGTTTATTACCGGAAATTTTGGGCACTAC 2339
DB 1387 GCGAAGCGCTTCTATCCCTACCCCAACCAAGTTTATTACAGAGAAATTTTTCGCTCATCAT 1446
QY 2340 GCGAAGCGGAAATTTGACACCCAGCATCATTTTTTTTCCAAAGTGGCTATGAAAAATGGC 2399
DB 1447 GGTAAAGCTTAATACGATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGGC 1506
QY 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAATCTGCCGTAACATC 2459
DB 1507 ATTAAGAGAAATTTAGGGCTTGAAGAGCAAGTGTGGCGGTAAAAAATTTGAGAAATATC 1566
QY 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAATACCGTTCGATCCGAAACAC 2519
DB 1567 ACTAAAAAGACATGCAATTCACGACCTACACGACCTACATGGAAGTCAATTTCTGAAACT 1626

```
QY 2520 TTCGAGGCTTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTG 2573
Db 1627 TACCATGTGTTGCTGGATGGCAAA-GAAGTAACCTCTAAACCCACCAATAAAGTG 1679

RESULT 15
US-10-282-122A-32950
; Sequence 32950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32950
; LENGTH: 2341
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-10-282-122A-32950

Query Match 25.5%; Score 734.6; DB 7; Length 2341;
Best Local Similarity 58.6%; Pred. No. 3.4e-157;
Matches 1408; Conservative 0; Mismatches 929; Indels 66; Gaps 5;

QY 207 TGAACCTCACACCCAAAGACGAGAAAGTTCTTGTATATATATGCGGGCGAAGTGCTA 266
Db 2 TGGAAATTAACACCAAGAGAAAAAGATAAAATTAATCTGCTTTTACTGCGAGGCTTTGTCAG 61

QY 267 GAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTG 326
Db 62 AAAGACGTTTGTAGCTTAAGGGATTAATACTTAATACCTGAAGCTGTGCGCTTGATTAGCT 121

QY 327 CCCATATTATGGACGAAGCGCGCTGGAAAAAACCCTTGCCACGCTTATGGAAGAGT 386
Db 122 GCGCCATTATGGAAGCGCAGAGAGG--GAAAAAGTTGCTCAATTAATGAGTGAAG 178

QY 387 GCATGCACTTTTGAAGAAAGATGAAGTAATGCGCGGGGTGGTAAATATGTTCCCGATC 446
Db 1258
```

```
Db 179 GACGTACTGTTTAAACCGCAGACCAAGTAATGGAAGGGGTGCCAGAGATGATAAAAGATG 238
QY 447 TAGGTGTAGAAGCCACCTTTCTGTATGGTATAGAAACTTGTAACTGTGAATGGGCCCATCG 506
Db 239 TTCAAGTATAGATGCACTTTCCCGATGGCAACCAAAATGGTTTCAATTCACCTCACTATTG 298
QY 507 AACCATGATGACACTTCAAGCGGGCGAAGTGAATTTGGTTGGATAAA--GACATCG 563
Db 299 TGT-AGGTAATAACATGATCCCGGTGAATTAAGAGTTAATGCGCATTTAGGCCATATTG 357
QY 564 AGCTCAATGACGCAAAAGAACTAACCGAACTTGAGGTTACTTAATGAAGGGGCTAAATCCT 623
Db 358 AACTGAATGCTGTCGCGAGACAAAACCATACAGGTGGCTAATCATGGCGATAGACCTG 417
QY 624 TGCATGTGGTAGCCATTTCCACTTTCTTTGAAGCTAAACAGGCACTAAAATTCGATCGTG 683
Db 418 TACAAGTCGGCTCTCATTTACCACTTTTATGAAGTTAATAGGCGCACTCAGGTTTGCACGAG 477
QY 684 AAAAGCCTATGGCAAAACGCTAGATATTCCTCTGGCGCAACACCGCTACGCAATGGGGCAG 743
Db 478 AAGAGACATTAGGTTTTCGTTTAAATATTCCTGCTGTAAGGCTGTGCGCTTCGAGCCCG 537
QY 744 GACAAACCCCAAGTGCAGTTGATTCCTCTTGGTGCAGTAAAAAAGTGAATGGCATGA 803
Db 538 GTCARAGCCGCACTGTTGAGTTAGTGCTTTTGCAGGAACAGTGAAATTTATGG----- 592
QY 804 ACGGGCTTGTGAATAACATCGCGGATGAACGCCATAAAACATAAAGCGCTTGACAAAGGCA 863
Db 593 -----TTTTTCATGGCAA 604
QY 864 AATCTACGGATTATCAAGTAAGGAGACTCCCATGAAATGAAACCAAGAAATATCTA 923
Db 605 AGTGATGGTAAATTTGGAGAGTGAGAAAAAATACTCTCACGTCACAGCTTATGCG 664
QY 924 AATACCTTACGACCCCAACCAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCA 983
Db 665 GATATGTTGGCCCAACACAGCGGATCGTTTGGATTAGCAGATACCGAGCTGTTCTT 724
QY 984 GAAGTAGAACATGACTATACCACTATGGGGAAGAACTTAAATTTGGCGGGTAAAAACT 1043
Db 725 GAAATTCGAAAAAGATTTCACCACTTATGGCGAAGAGGTCAAATTTGGTGGTAAAGTT 784
QY 1044 ATCGGTGAGGGTATGGGTGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATC 1103
Db 785 ATTGCTGATGTTGGGGCAAGCCAAAGTTGTTAGTGTGCTGATGTGTCGATGTTCTGATC 844
QY 1104 ACTAACCGCATGATTATCGACTACACCGGGATTTCACAAAGCCGACATTTGGGATTAATAAC 1163
Db 845 ACCAATGCCATTATTTAGATTATTTGGGCAATTGTAAAGCAGATATTTGGCATTAAGAT 904
QY 1164 GGCAAAATCCATGGCATTGGCAAGCGAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCT 1223
Db 905 GGCGTATTGTCGGTATTGGCAAGCGAGGTAATCCAGATGTTTCAGCCCAATGT-----G 958
QY 1224 CATATGCTGTGGTGTGGGCAAGAACATAGCAGGGGAAGGTATGATTAATTACCGCT 1283
Db 959 GATATTGTCATTGCCCCCGGAAACAGAAAGTTGTGCGGGAGAGGTTAAATATAGTCTGCT 1018
QY 1284 GGGGGAATCGATTACACACACCCACTTCTTTCTCCACAACTTCCTTACCGCTCTAGCC 1343
Db 1019 GGTGGTATTGATACCCATATCCACTTTATTTGTCCACAAACAGCCCAAGAGGTCTCGTT 1078
QY 1344 AATGCGGTACAACCATGTTTGGAGGGCGGCAAGGTCTCTGTAGATGGCACGAATCGACT 1403
Db 1079 TCTGCGGTAAACCACTTTATTTGGTGGAGAACAGGCCCTGTGGCAGGTACTTAATGCAACC 1138
QY 1404 ACTATCACTCGGGCAAAATGGAACCTTGACCGCATGTTGGCGGCAAGCAGAGATTTCT 1463
Db 1139 ACAGTTACCCCGGTATTTTGGAAATATGTACCGCATGTTAGAGCGGCTGGATGAATACCT 1198
QY 1464 ATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACTGTAGAACA 1523
Db 1199 ATTAATGTGGGGTTATTTTGGTAAAGGTTGTGTGCTAGTCAGCCGGAAGCAATCCGCGAACA 1258
```

QY	1524	GTAGAGCGGCGCGATTCGGTTTAAATTCATGAAGACTGGGGCACAACACCAAGTGGC	1583
Db	1259	ATAACAGCGGGTCTATAGGTCTTAAATATACATGAAGACTGGGGGCAACGCCAATGGCA	1318
QY	1584	ATCGATCAGTCTGTGAGCGTGGCAGATGAATAGGATGTCGAAGTTTGTATCCACCGAT	1643
Db	1319	ATTACAATTTGCTTAATGTCGCGGATGAATGGATGTACAGGTGGCTATTCACTCTGAC	1378
QY	1644	ACAGTCAATGAGCGAGTTATGTAGATGACACCCTAAATGAATGAACGGGGCGGCATC	1703
Db	1379	ACCTTAAATGAAGGTGTTTTATGAAGAGACAGTAAAGCCATTGCCGTCGAGTGATC	1438
QY	1704	CATGCCCTACCAATTTGAGGAGCGGGTGGAGACACTCACCTGATGTATTCAACATGGCA	1763
Db	1439	CATGTATTCATACCGAAGCGCAGGTGGTGTGTCATGCCCTGATGTGATCAAGTCGGTA	1498
QY	1764	GGGAGCTCAATATTTCTACCTCTCCACACCCCACTATTCCCTATACCATTAATACG	1823
Db	1499	GGAGAGCCCAATATTTTACCTGCTCAACCAACCCCAACGATGCTTTATACCATTAATAC	1558
QY	1824	GTTCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAG	1883
Db	1559	GTGGACGAGCATCTTGATATGTGTGATGTGTCTCATCTCGATCCCTCTATTCTCTGAA	1618
QY	1884	GATTTACAATTTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTC	1943
Db	1619	GATGTGGCATTTGCTGNACTCTGTAATCTGTGCGGAACCATTTGCTGCAGAGATATCTTA	1678
QY	1944	CATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGGTGCGGC	2003
Db	1679	CATGATATGGGGCAATTTCCGTGTGATGTCGTGAGACTCAACAGCCATGGGAGTCGGA	1738
QY	2004	GAAATGATTCCTCGAATTTGGCAGACTGCGGATAGAATAAAGAAATTTGGTAAGCTT	2063
Db	1739	GAAATGATTCCTAGCAGCTTTGGCAGTGTGCACATAAATGAATTTGCAACGAGGCACTTA	1798
QY	2064	CCTGAAGATGGCAAGACATACGATAATTTCCGCAATTAAGCGTACATCTCCAAATACACT	2123
Db	1799	GCGGTGATAGCGCAGATATGATTAATATCGTATTTAAAGTTATTCGTTAATACAG	1858
QY	2124	ATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGGGCAAG	2183
Db	1859	ATTAACCCCGCACTGGCACAATGGTATGCTCATACGGTGGATCAATAGAAAAGGTAAA	1918
QY	2184	ATCGCCGACTTGGTGTGTGGAATCTGCTTTTGGCGGTAAACCCCAAAATCGTGATC	2243
Db	1919	CTTGGCGATATCGTGTATGGGATCTGCTTTCTTTGGCGTCAAAACCGGCACCTTATCATA	1978
QY	2244	AAAGCGGTATGGTGTCTTCTGNAATGGGGATTTCTAACGCGTCTGTGCCACTCCC	2303
Db	1979	AAAGTGGTATGGTGTGCTTATGCGCAATGGGGGATTAATATGCGGCTATTCAACACCG	2038
QY	2304	CAACCGTTTATTACCGGAAATGTTGGGCATCACGCAAGCGGAAAATTGACACACG	2363
Db	2039	CAACCGTTTCAATATGCTCAATGATGCTGTTAGGAAAAGCCAAATATCAACGTCG	2098
QY	2364	ATCACTTTTGTTCAAAGTTCGCTTATGAAAATGGCGTGAAAAGAAAGCTGGGCTTAGAG	2423
Db	2099	ATGATCTTTATGTCAAAAGCGGTATTGAGCGGGAGTCCAGAAAATTAGGCTTAAAA	2158
QY	2424	CGCCAAAGTTCTACCGTCAAAAACCTCGGTAACATCACCAAGAAAGCTTCAAGTTCAAC	2483
Db	2159	AGCTTAAATTTGGTGTGTGGAGGCTGTGCTCATATCACAAGAGCTTCGATGATCCCAAT	2218
QY	2484	GACAAAACGGCAAAAATCACCGTCGATCGAAGAACTTCGAGGTCTTTGTAGATGCAAA	2543
Db	2219	AACATATGTTCTCATATCGAATTTAGATCCAAAACCTTACATTGTTAAAGGGATGTGTA	2278
QY	2544	CTCTGCACTCTAAACCCCACTCGCAAGTGCCTCTAGCCCGAGCGCTTCTTCTTAG	2603
Db	2279	CCACTGGTTTGTGAGCGACGACTGAATTACCGATGGCTCAACGCTATTCTTTATTTAA	2338

QY	2604	GCA	2606
Db	2339	CCA	2341

Search completed: November 29, 2005, 08:04:19
Job time : 2060 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:34:45 ; Search time 563 Seconds
(without alignments)
760.989 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgragatttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3205263 seqs, 74304013 residues

Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	612	21.2	3164	1	US-10-793-626-3828
2	612	21.2	3234	1	US-10-793-626-3853
3	578.6	20.1	1716	1	US-10-793-626-117
4	455	15.8	1612	1	US-10-485-517-26
5	424.4	14.7	3475	1	US-10-793-626-4032
C 6	216.2	7.5	526	1	US-10-834-151-2
C 7	97.4	3.4	309	1	US-10-793-626-121
8	74.2	2.6	336	1	US-10-793-626-115
9	46	1.6	2259	1	US-10-793-626-4078
C 10	40	1.4	3373	1	US-10-793-626-4389
C 11	40	1.4	3927	1	US-10-793-626-3926
C 12	39.8	1.4	3923	1	US-10-793-626-4109
C 13	37.8	1.3	1809	1	US-10-793-626-2361
C 14	37.8	1.3	3591	1	US-10-793-626-4033
C 15	37.2	1.3	2988	1	US-10-793-626-4132
16	37.2	1.3	4210	1	US-10-793-626-3761
17	37	1.3	3362	1	US-10-793-626-3731
C 18	37	1.3	4041	1	US-10-793-626-3689
19	36.2	1.3	738	1	US-10-793-626-2909
20	36.2	1.3	3218	1	US-10-793-626-4229
C 21	36.2	1.3	3997	1	US-10-793-626-4220
22	35.8	1.2	14941	1	US-10-821-234-771
C 23	35.6	1.2	1017	1	US-10-793-626-1003

C 24	35.6	1.2	1439	1	US-10-485-517-41	Sequence 41, Appl
C 25	35.6	1.2	1500	1	US-10-485-517-56	Sequence 56, Appl
C 26	35.6	1.2	3189	1	US-10-793-626-3694	Sequence 3694, Ap
C 27	35.6	1.2	3223	1	US-10-793-626-4222	Sequence 4222, Ap
C 28	35.6	1.2	3285	1	US-10-793-626-4141	Sequence 4141, Ap
29	35.6	1.2	340000	7	US-11-102-978-3	Sequence 3, Appli
C 30	35.4	1.2	3609	1	US-10-793-626-3919	Sequence 3919, Ap
C 31	35.2	1.2	4248	1	US-10-793-626-3885	Sequence 3885, Ap
C 32	35	1.2	2308	1	US-10-510-386-61	Sequence 61, Appl
C 33	34.6	1.2	3473	1	US-10-793-626-3343	Sequence 3343, Ap
C 34	34.4	1.2	1274	1	US-10-927-641-34	Sequence 34, Appl
C 35	34.4	1.2	1299	1	US-10-793-626-779	Sequence 779, App
36	34.4	1.2	3037	1	US-10-793-626-3904	Sequence 3904, Ap
37	34.4	1.2	3073	1	US-10-793-626-3606	Sequence 3606, Ap
38	34.4	1.2	3119	1	US-10-793-626-4270	Sequence 4270, Ap
39	34.4	1.2	3536	1	US-10-793-626-3957	Sequence 3957, Ap
40	34.4	1.2	3645	1	US-10-793-626-4110	Sequence 4110, Ap
41	34.4	1.2	3854	1	US-10-499-715-5	Sequence 5, Appli
42	34.4	1.2	3932	1	US-10-793-626-3420	Sequence 3420, Ap
43	34.4	1.2	4020	1	US-10-793-626-3434	Sequence 3434, Ap
C 44	34.2	1.2	3473	1	US-10-793-626-3343	Sequence 3343, Ap
C 45	34.2	1.2	3546	1	US-10-793-626-3457	Sequence 3457, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-3828
; Sequence 3828, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3828
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3828

Query Match	21.2%	Score 612;	DB 1;	Length 3164;
Best Local Similarity	57.4%	Pred. No. 5.2e-139;	Mismatches 845;	Indels 52;
Matches 1211;	Conservative 0;			Gaps 4;
QY	543	TTGGTTGGCATAAGACATCGAGCTCAATCGAGCGCAAGAGAGTAACCGAACTTTGAGTTA	602	
Db	468	TTGTTAAAAATCTAGTAATAGAGTCAATAACATCATCCCGAAACGGTGTGAAGTAA	527	
QY	603	CTAATGAAGGGCCTAAATCCTTTGCATGGGGTAGGAGTTCACATTCCTTTGAAGTACA	662	
Db	528	AAATACAGGCGATAGACCTATACAAAGTAGGTTCACATTCCTTTTCGAAGCAATA	587	
QY	663	AGGCACTAAATTCATCGTGAAGAGCCTATGGCAAGCGCTAGATATTCCTCTGTGGCA	722	
Db	588	AAGCATTAAGAAATTTGATCGTGAGAAAGCATATGTAACATTTGATATTCCTCGAGGAG	647	
QY	723	ACACCTACGATTTGGGCGAGCAACACCCGAAAGTGCAGTTGATTCCTCTGTGGCA	782	
Db	648	CTCAGTCAGATTGGAACCTGGAGATGAAAAAGTACAACTTTCGAATATTCCTGAC	707	
QY	783	GTAAGAAAGTATGGCATGAACGGGCTTGTGAATTAAC---	839	
Db	708	GACGTAAATTTATGGATTCCGTGGTTAGTCGATGGCGATATTGACGAAGAACCGGTAT	767	


```

Qy 840 AACATAAGCGCTTGACAGGCGAAATCTCAGGATTTATCAAGTAAGAGACTCCCATG 899
Db 768 TCGGTCAATATGATTCAAATCAAAACGCGCGTTAAAGATGACAGCGAAGCATG 827
Qy 900 -----AAAAAGAAAAAAGAAATATGTA 925
Db 828 CGAATAAAAAAGTGTAAATAGCATGATGTTAAATGACACAAATCTCAATCACAAAG 887
Qy 926 TACCTAGGAGCCCAAGGCGATTAAGTGGCTTGAGGATACCGCATCTTTGGGCGA 985
Db 888 TCTTTATGACCAACTGAGAGACTCTGTGAGATTAAGAGATACGAATCTGTTTGACA 947
Qy 986 AGTAGAATGATCTATACCACTTAAGCGAAGAACTTAATTTGGCGGCGGTAATAT 1045
Db 948 AGTTGAAAAAGACTATGCAAAATTAAGAGATGAGACTCTTCGGTGGCGAAAAATCAT 1007
Qy 1046 CCGTAGAGGTATGGGTGCA-----GCAATAGCCCTGATGMAAACACCTTAGATT 1096
Db 1008 TCGGATGTATGGCTCAAAATCTTAATGTGACAAAGATGATATAAAATGTAGCCGATT 1067
Qy 1097 AGTCATCACTAAGCGATGATTAATGACTACACCGGATTTACAAAGCGCATTTGGAT 1156
Db 1068 AGTTTAACTAACCGCATTAATTAATGATTAATGACAAAGATGTTAAAGCATATCGAT 1127
Qy 1157 TAAAAAGCGCAAAATCCATGCGATTGGCAGAGCAAGAAACAGACATGCAAGATGGGT 1216
Db 1128 TAAAAATGTATATTTTAAATGATGCTGTAAGCTGMAAACCAATATTAATGATTAAGT 1187
Qy 1217 AAGCCATATATGTGCTGAGGTGTGGGACAGAAAGCACTAGCAGGGAAGGTATGAT 1276
Db 1188 -----TGACATCATCATTTGGTGCAACATGATATATGCTGTGTAAGGTAATTTGT 1241
Qy 1277 TACCGCTGGGGGAATTCGATTTACACACCTTCTTCTCACAACAAATTCCTACCGC 1336
Db 1242 TACTGCGGCGGTATCGATACACAGTCACTTCAATCCTGGAACAGCTGAAGTTGC 1301
Qy 1337 TCTAGCCAAATGGCGTTACCAACATGTTTGAAGGCGGCAAGGCTGTAGATGGCAGCA 1396
Db 1302 ACTTGAAGTGTATTAACAGCATATCGTGAAGAACTGGTCTTGAAGGTCTAA 1361
Qy 1397 TGGCACTACTATCACTCCGCGCAATGGAATTGCAACCGATGTTCCGCGCAGCAGAGA 1456
Db 1362 AGGCATCTACTGTAAACACAGGACCTTGGATATTCATGCAATGTTAAGAGCAGGAGA 1421
Qy 1457 GTATTTCTATGATGTGGGCTTTTGGGCAAGGCAATAGCTTACAAAAAACAATTGT 1516
Db 1422 GATGCTTATTAATGATGATTTTACTGTAAAGGTCAAGCTGTCAATCACTGCACTTAT 1481
Qy 1517 AGAACAAGTAGAAGGCGGCGGATTTGGTTTAAATTTGATGAAGACTGGGCGCAACACC 1576
Db 1482 TGAACAAATTCATGAGGCGGTATAGGCTTTAAAGTACAGTAAGATGGGAGCTACACC 1541
Qy 1577 AAGTGCATGCATCACTGCTTGAGCGTGCGAGATGAATACATGTGCAAGTTTGTATCCA 1636
Db 1542 TTTACGATTAAGTATGATTAAGATGAGTTCAGATGAGTTGATGTTCAAGTGGCTTACA 1601
Qy 1637 CACCGATCACTCAATGAGCGAGTTATGTAGATGACACCTTAAATGCAATGACGCGCG 1696
Db 1602 TGCAGATCACTTAATGAACTGATTTATGMAATACAAAGGCTGCTGTAAAGATG 1661
Qy 1697 CGGCATGCATGCTTACCATGAGAGGAGCGGTGAGAGCACTACCGCATGATTATGAC 1756
Db 1662 TGTATGTATATGTATCACTAGAAAGAGCTGTGTGTCTATGACCTGATTAATCAA 1721
Qy 1757 CATGCGAGCGAGCTCAATATTTCTACCTCTCCACACCCCACTATTCCTATACAT 1816
Db 1722 ATCAGCTGCATATTCAAACATCTTACCTTCTTCAAAACCAATTAACCTTACATCA 1781
Qy 1817 TAAATCGGTGCAAACTTGAACATGTCTATGACATGACCAACCACTTACAGCAACGAT 1876
Db 1782 CAACACTGTATGATGAACATTTAGACATGGTATGATTACTACACCACTTAAATGCTTCAAT 1841

```

```

Qy 1877 CCGGAGGATTTACATTTTCTCAAAAGCCGTATCCGCCCGGCTTATCGGCGTGAAGA 1936
Db 1842 ACCAGAAAGACATTTGATTTGAGATTTCTGTATACGTAAAGAAACTATAGACGAGAA 1901
Qy 1937 TGTGCTCATATATATGAGGTGATATGCGATGACAAAGCTCGATTTGCGAAGCAATGGGCG 1996
Db 1902 CGTATTAAGATATGAGGCGTATTTAGTATGTTAGTAAAGTTCAAGTTACAAAGCAATGGAC 1961
Qy 1997 TGCAGGCGAATGATTTCTCGAACTTTGGCAGACTGCGATTAAGAAATTAAGAAATTTGG 2056
Db 1962 TGTGCTGAAGTTGTAACACGTAATCTTGCCAGATTTGACACCGTATGAAGAAACACGCG 2021
Qy 2057 TAACTTCTGAGAGTGGCAAGATACGATATTTCCGATTAAGCCGTACATCTCAA 2116
Db 2022 ACCATTAAGATGATCTTTGAATATACAGATTAATGTATGATTAACCTTACATGCAAA 2081
Qy 2117 ATACATTAACACCCGCTTTGACCAAGCGGTGAGGAGTATCGGCTGTGAGAA 2176
Db 2082 ATATCAATCAATCTGCAATTAACATGATTTTCTGATTAATGTTGATCTGTAAGAC 2141
Qy 2177 GGGCAAGATCGCGCACTGTGTGTGGAATCTGCTTTTGGCGTAAACCCAAAT 2236
Db 2142 GGTAAACTTCCGATTTAGTATGTGGAACAGAAATTTCTGCGTGCACAAACCGATCT 2201
Qy 2237 CGTATCAAGCGGTATGTGTCTTCTGAAATGGCGATTTAAAGCGTCTGTGTC 2296
Db 2202 TGTGTTAAAGTGCATGATTAATCTCAGCAGTAATGGTATGCTAAATGGCTCCATACC 2261
Qy 2297 CACTCCCAACCGGTTTATACCGGAAATGTTTGGGATCAACGCGCAAGCGAAATTTGA 2356
Db 2262 AACATCAGAGCTTTGAAATATCGCAAAATGATGTGCAATTTGGTGTACATTACACA 2321
Qy 2357 CACGACATCACTTTGTTTCCAAAGTGCCTATGAAATGCGTGAAGAAAGAAAGCTGGG 2416
Db 2322 TACTGCTATGATTTTGTGTTCTTACACCTGCAATATGAAGATTTATTCGCACTCAA 2381
Qy 2417 CTTAGAGCGCAAGTTCTACCGGTCAAAATCGCGTAAATCATCACCAAGAAAGCTTCAA 2476
Db 2382 TCTAAGAAAGATGTTTGCAGCTTGAAGAAATTTGAAGAAATTTAACTAAGGCGATATGA 2441
Qy 2477 GTTCAACGCAAAACGCGAAATATCAACGTCGATTCGAAACCTTGGAGGTCTTTGTAGA 2536
Db 2442 AATATATATGCTTACACCTTAAATATGATGATCCACAAACATGAGGTATCGTTGA 2501
Qy 2537 TGGCAAACTGTGCACTCTTAAACCAACCTGGAAGTCTCTAGCCACGCTCACTTT 2596
Db 2502 TGTATTAATATCAAGATGAAGACAGCAAGAAATTAACATTAACCAAGATTACTTCTT 2561
Qy 2597 CTTCTAGG 2604
Db 2562 ATTCTAGG 2569

```

RESULT 2
 US-10-793-3853
 ; Sequence 3853, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; PRIORITY FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 3853
 ; LENGTH: 3234
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-3853
OTHER INFORMATION: nucleic acid sequence

Query Match 21.2%; Score 612; DB 1; Length 3234;
Best Local Similarity 57.4%; Pred. No. 5.3e-139;
Matches 1211; Conservative 0; Mismatches 845; Indels 5

Qy	543	TTGGTTGCGATAAAGACATCTCAGAGCTCAATGCGAGGCAAAAGAGTAACCGAACTTGAGGTTA	602
Db	537	TTGTTAAAAATACCTGAATAGAAGTCAATAAACATCATCCCGAAACGGTGATTTGAAGTAA	596
Qy	603	CTAATGAAGGGCCTAAATCCCTTGATGATGGGTAGCCATTTCCACCTTCCTTCGAAGCTAAC	662
Db	597	AAAATACAGGCGATAGACCTATACAGTAGGTTCCATTTCCATTTTTCGAGCCAAATA	656
Qy	663	AGGCACTAAAAATTCGATCGTGAAAAAGCCTATGCGCAAAACGCCCTAGATATTCCTCTGSCA	722
Db	657	AAGCATTAGAATTTGATCGTGAGAAAAGCATATGGTTAAACATTTTGGATATTCCTGCGAGG	716
Qy	723	ACACGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGCCA	782
Db	717	CTGCAGTGAGATTTGAACCTCGAGATGAAAAAAGTACAACTTGTTCGAATATTCCTGGAC	776
Qy	783	GTAAAAAGTGAITGGCATGAACGGGCTTGTGAATTAAC---ATCGCGGATGAAGCCCAT	839
Db	777	GACGTAAAAATTTATGGATTCCGTGCTGTTTGTAGTCGATGGCGATATTGACGAAGAACGCTAT	836
Qy	840	AACATAAAGCCCTTGACAAAGCGGAAATCTCACGGATTTATCAAGTAAGAGACTCCCATG	899
Db	837	TCCGTCCAATATGATTCAAATCAAACCGCGCGCTTAAATAACGATGTCAGCGGAAGACAAATG	896
Qy	900	-----AAAATGAAAAAACAAGATATGTAAA	925
Db	897	CGAATAAAAAAGGTGTAATAAGCATGAGTTTTAAATATGACAACTCTCAATACACAAG	956
Qy	926	TACCTACGACCCACAAAGCGGATAAAGTGGCTTAGAGATACCGATCTTTGGGGCAGA	985
Db	957	TCCTTATGCGACCAACTGTAGGAGACTCTGTGAGATTTAGGAGATACGAACCTTTGTTGACA	1016
Qy	986	AGTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTTCGCGCGGTAAACACTAT	1045
Db	1017	AGTTGAAAAAGACTATGCAATTTATGAGATGAAGCTACTTTTCGTTGCGGGAATAATCAAT	1076
Qy	1046	CCGTGAGGGTATGGGTGAGA-----GCAATAGCCCTCATGAAAAACCCCTAGATTT	1096
Db	1077	TCGTGATGTTATGCTCAAAATCCTAATCTGCAAGAGATGATAAAATGTAGCCGATTT	1136
Qy	1097	AGTCATCACTAAACGGATGATTATCGACTACACCGGATTTTACAAGCCGACATTTGGAT	1156
Db	1137	AGTTTTAACTAACCGCATTAATTTATGATTTATGACAAGATTTGTTAAAGCAGATATCGGAAT	1196
Qy	1157	TAAAAACGGCAAAATCCATGCGATTTGGCAGCGAGAAAAAAGACATGCAAGATGGCGT	1216
Db	1197	TAAAAATGTTATATTTTAAGATCGGTAAAGCTGGAAACCCAGATAAATGGATAACGT	1256
Qy	1217	AAGCCCTCATATGTTGTTGGGTGGGCAAGAGACACTAGACGGGGAAGTATGATTAT	1276
Db	1257	-----TGACATCATATGGTGCAACAACCTGATATTATTCTGCTGTAAGGTAAATTTGT	1310
Qy	1277	TACCGCTGGGGGAATCGAATTCACACACCCACTTCCTTTCTCCACAAACAATTTCCCTACCGC	1336
Db	1311	TACTGCGCGGTATCGATACACACGTGCACTTCATCAATCTGNAACAAGCTGAAGTTGC	1370
Qy	1337	TCTAGCCATGGCGGTACAAACCATGTTTGGAGCGGCAACAGTCTCTGTAGATGGCAGCAA	1396
Db	1371	ACTTGAGAGTGGTATTAACAACGCATATCGTGTGAGGAACCTGGTGTCTCTGAAGGTGCTAA	1430
Qy	1397	TGGCACTACTATCACTCGGCGAAATGGAATCTCGACCGCATGTTTCGCGCAGCAGAGA	1456
Db	1431	AGCGACTACTGTAAACACAGGACCTTTGGCATATTCATTCGATGTTTAGACGACGACAGA	1490
Qy	1457	GTATTTCTATGAATGTGGGCTTTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACCTGT	1516

1491	Db	GATCCCTTAATAAGTGAAGATTACTGGTAAAGGTCAAGCTGTGCAATCATACTGCACCTTAT	1551
1517	Qy	AGAAACAAGTAGAAGCGGCGCGAATTGGTTTTAAATTGTCATGAAGACTGGGGCACAACACC	1576
1551	Db		1610
1577	Qy	TGAACAAATTCATGCGAGCGCTATAGGCTCTTAAAGTACATGAAGATTGGGGAGCTACACC	1636
1611	Db		1670
1637	Qy	AAGTGGCATCGATCATCTGCTTGAAGCGTGGGAGATGAATACGATGTGCAAGTTTGTATCCA	1696
1671	Db		1730
1697	Qy	TTGAGCATTAAGTCATGCATTAGACGTTTAGACGTTAGATGACACCCCTAAATGCAATGAACGGCG	1756
1731	Db		1790
1757	Qy	TGATTGTCATATGATCATACTGAAGGAGCTGGTGGTGGTCATGCACCTGACTTAATCAA	1816
1791	Db		1850
1817	Qy	CATGGCAGGCGAGCTCAATATTCTACCTCTCCACACCCCACTATTTCCTTATACCAT	1876
1851	Db		1910
1877	Qy	ATCAGCTGTCATATTCAAACACTCTTACCCTCTCTCAAAACCCCAACATTCCTTACACTCA	1936
1911	Db		1970
1937	Qy	TAATACCGTTGTCAGAACACTTAGACATGCTCATGACATGCGCACCACTAGACAAACGCAT	1996
1971	Db		2030
1997	Qy	CGTATTACAAGATATGGCGCTATTTAGTATGTTAAAGTTTCAGATTTCACAGCAATGGGACG	2056
2031	Db		2090
2057	Qy	TGCAGGCGAAGTGATTCCTCGAATCTGGCAGACTGCGGATAGCAATAAATAAAGAAATTTGG	2116
2091	Db		2150
2117	Qy	TGTCGGTGAAGTTGTAACACGTAATCTGGCAAGTTGCACACCGTATGAAGAGCAACCGCG	2176
2151	Db		2210
2177	Qy	ATATACAAATCAATCCTGCCATTACACATGTAATTTCTGACTATGTTGGATCTGTAGAACG	2236
2211	Db		2270
2237	Qy	GGGCAAGATCGCGACTTGGTGTGTGGATCCTGCCTTTTTTGGCGTAAACCCCAAAAT	2296
2271	Db		2330
2297	Qy	CGTGATCAAAAGCGGTATGTGTGCTTCTCTGAAATGGGCGATTCTAAACGCTCTGTGCC	2356
2331	Db		2390
2357	Qy	TGTTGTTAAAGTTGGCATGATTAATCTCAGCAGTAAATGGTGATGCTAAATGGCTCCATACC	2416
2391	Db		2450
2417	Qy	CATCCCCAAACCGGTTTATTAACCGGAAATGTTTGGGCATACCGCAAGCGGAAATTTGA	2476
2451	Db		2510
2477	Qy	AACATCAGACCTTTGAAAATATCGCAAAATGTTATGTGTCATATTTGGTGGTAACTATACACA	2536
2511	Db		2570
2537	Qy	CTTTAGAGCGCAAGTTCTACCGGTCAAAACTGCCGTAAACATCAACCAAGAAAGACTTCAA	2596
2571	Db		2630

QY 2597 CTCTAGG 2604
Db 2631 ATTCTAG 2638

RESULT 3

US-10-793-626-117
; Sequence 117, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-117

Query Match 20.1%; Score 578.6; DB 1; Length 1716;
Best Local Similarity 59.5%; Pred. No. 4.8e-131;
Matches 1021; Conservative 0; Mismatches 679; Indels 15; Gaps 2;

QY 898 TGAATAATCAAAAAACAAGATATGTAAATACCTACGACCAACCAAGCCATAAGTGC 957
Db 8 TTTAAATGACACAACTCTCAATACCAAGCTTTATGACCACTGTAGAGACTCTGTGA 67
QY 958 GCTTAGAGATACCATCTTTGGGCAAGTAAACATGATCATATACCACTATGCGCAAG 1017
Db 68 GATTAGAGATACCAACTCTTTGGGCAAGTAAACATGATCATATACCACTATGCGCAAG 127
QY 1018 AACTTAATTTGGCGCGGGTAAACTATCCGTGAGGGTATGGGTGAGA-----GCA 1068
Db 128 AAGCTACTTTCGTGGCGCGGAATAATCAATTCGTGATGATGGCTCAAAATCCTAATGTGA 187
QY 1069 ATAGCCCGATGAAACACCTTAAGTTAGTCACTAAGCGCATGATATCGACTACA 1128
Db 188 CAAGAGATGATAAATAATGTAGCCGATTTAGTTTAACGATTAATTAATGATTATG 247
QY 1129 CCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGCGAAG 1188
Db 248 ACAAGATTGTTAAAGCAAGATTCGGAATTTAAAAATGGTTATTTTAAAGTCGTTAAG 307
QY 1189 CAGAAACAGAGACATGCAAGATGCGTAAAGCCCTCATATGGTCTGTGGGTGCGGCAAG 1248
Db 308 CTGAAACCCAGATATATATGATTAAGT-----TGAATCATCATTTGGTGTCAACAAC 361
QY 1249 AAGCATGACGAGGGAAGTATGATTTAACCGCTGGGGGAATCGATTACACACCAC 1308
Db 362 ATATATATGCTGCTGAAGGTAAATTTGTTACTGCGGCGGTATCAATACACAGTGCAC 421
QY 1309 TCCCTTCCCAACAATTCCTACCGCTCTAGCCATTTGCGCAATGGCTTCAACATGTTTGA 1368
Db 422 TCATATATCTGGAACAAGCTGAAGTTGCACTTGAAGTGTATTTACAGCGCATTCGGTG 481
QY 1369 GCGGACAGTCTCTGTAGATGCAAGATGCGATCTACTATCACTCCGGCAATGGAAC 1428
Db 482 GAGGAATCTGTCTCTGAAAGTGTAAAGCACTACTGTAAACACAGGACCTTGGCATA 541
QY 1429 TGCACCGATGTTGGCGCGGACAGAAAGTATTTATGATGTTGGGCTTTTGGGCAAG 1488
Db 542 TTCATCGATGTTGAAGACAGCAAGAAAGATGCTATTAATGATGAGATTACTGTGAAG 601

QY 1489 GCAATAGCTTAGCAAAAAACAATTGTAGAAACAAGTGAAGACGGGCGCGATGGTTTA 1548
Db 602 GTCAAGCTGTCAATCATCTGACATTAATTAAGAAACAATTCATGACGGCCCTATAGTCTTA 661
QY 1549 AATTGCAATGAAGTCTGGGGCAACAACCAAGTGGATGATCACTGCTTGAAGCGGAG 1608
Db 662 AAGTACATGAAGATTTGGGAGCTTACCTTGAAGCTTAAGTATGATTTAGAGTTGAG 721
QY 1609 ATGAATACGATGTGCAATTTGTATTCACACCGATACAGTCAATGAGGAGTTATGAG 1668
Db 722 ATGAGTTGATGTTCAATCGCTTTACATGCAAGATACATTAATGAACGTGATTTAG 781
QY 1669 ATGACACCTTAATGCAATGAAGCGGCGCGCATCATGCTTACCACTTGAAGGAGCG 1728
Db 782 AAGATACAAATGCTGCTGTGAAGATCGTGAATTCATATGATATCATCTGAAGAGCTG 841
QY 1729 GTGAGGACATCTACCGTATGTTATTCACATGCGGAGGAGGAGTCAATTCACCGCTCT 1788
Db 842 GTGGGTGATGACACTTACCTTAATCAATCAGCTGCAATTCAAACATCTTACCTTT 901
QY 1789 CCACACCCCACTATTCCTATATACATTAATATACGGTTGACAGACCTTAGACATGCTCA 1848
Db 902 CTACAAACCAACATTAACCTTACACTCAACAACATGATGATGATTAAGATGATTA 961
QY 1849 TGACATGCCACCACTAGCAAAACGCAATCCGCGAGATTTACATTTTCTCAAGCCGTA 1908
Db 962 TGATTAATCACTATTAATGCTTCAATACCAAGAGATTCATGCTTACAGATTCCTGTA 1021
QY 1909 TCCGCCCGGCTCTATCCGCTGTAAGATGTGCTCATGATATGGGTGATCGGATCA 1968
Db 1022 TACGTAAAGAACTATACAGACAGAAAGAGTATTAACAAGATATGGCGATTTAGTATG 1081
QY 1969 CAAGCTCGAATTCGCAAGCAATGGGCGGTGACAGCGAAGTATTCCTGAACTTGGCAGA 2028
Db 1082 TAGATTCAATTCACAGCAATGAGACGTGTGCGTGAAGTTTAAACGATCTTGGCAAG 1141
QY 2029 CTGCGATTAAGATTAATAAAGATTTGGTAAAGCTTCTGAAGATGCAAAAGATTA 2088
Db 1142 TTGCAACCGATTAAGAAAGAACACCGGACCTTATGATGTGATCTTGAATATCAAGATA 1201
QY 2089 AATTCCGATTAAGGCTTACATCTCAATATACATCTATCAACCCGCTTGAACCCAGCG 2148
Db 1202 ATATATGATTAATAAGTTACATTTGCAAAATATACATCATCTCCCATTAACATGTA 1261
QY 2149 TGAAGGATATATGGCTCTGTGAAAGAGGCAAGATGCGGACTTGGTGTGGAATC 2208
Db 1262 TTTCTGATATGTTGATGATCTGTAAGAGCGGTAACCTTGCATTTAGATGTGGAAC 1321
QY 2209 CTGCTTTTGGCGTAAACCCCAAAATCGTATCAAAAGCGGTATGTGATCTTCTCTG 2268
Db 1322 CAGAAATTCCTGGGTGCAAAACCGAATCTTGTATTAAGGTGGCATGATTAATCAGCAG 1381
QY 2269 AAATGGGCGATTTAAACCGCTCTGTGCCCATCTCCCAACCGGTTATTAACCGGAAATG 2328
Db 1382 TAAATGTGATGCTATGCTCATACCAACATCAGAGCTTTGAAATATCCAAATAT 1441
QY 2329 TTTGGGATACGAGCAAGGGAATTTGACACAGATCACTTTGTTTCAAAAGTGCCT 2388
Db 1442 ATGTCTAATTTGGGTGTAACATTAACATATCTGATATGATCTTTGTTTCAACCTGAT 1501
QY 2389 ATGAATAATGGCTGAAGAAAGAGTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAACT 2448
Db 1502 ATGAATAATGGATTTATATCTCACTCAATCTTAAAGATATGATGAGCAAGTTAGAAATA 1561
QY 2449 GCGGTATCATCAACAGAAAGCTTCAAGTTCAACGACAAACGCGAAATATCAACGTG 2508
Db 1562 TTAGAAATTTAATAAGGACATATGAAGAAATATATATCTACCTTAAATATGATGTAG 1621
QY 2509 ATCCGAAACCTTCAGGCTTTTGAAGATGCGAAACTGCACTCTTAACCAACCTCG 2568
Db 1622 ATTCACAAACATATGAGATGATGATGATGATGATTAATAATCAACAGTGAAGCAGAAC 1681
QY 2569 AAGTGCTCTAGCCGACGCTACACTTTCTTAG 2603

QY 774 TTGGTGCACGTAAGGATGTTGGCATGAACGGGCTGTGTAATPACATCGCGATGAA 832
DB 185 ATTCTGACGACGTAAGATTAATGATTCGCTGGTTTATGATGCGCATATTGACGAA 243

RESULT 9

US-10-793-626-4078
; Sequence 4078, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4078
; LENGTH: 2259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4078

Query Match 1.6%; Score 46; DB 1; Length 2259;
Best Local Similarity 48.2%; Pred. No. 0.061;
Matches 158; Conservative 1; Mismatches 166; Indels 3; Gaps 1;

QY 7 ATTTCCARCACTTCAAGCACATATGATCCGTGTGGGTGGTAAATTCRACCTTGT 66
DB 1907 AATTTCACTCAATCAATTAATTTTACTAAATATGTTTATGACGTGATATTTT 1966
QY 67 TAATRCATTTATTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 126
DB 1967 CATATTTATTTCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2026
QY 127 ATATTAAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 186
DB 2027 CAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2086
QY 187 TTAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 243
DB 2087 TAAATGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2146
QY 244 ATATTAATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303
DB 2147 AGTTGTAGCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2206
QY 304 CGAAGCCATTGCTTACATTTAGTCCCAT 331
DB 2207 TGAAGCACTGCTTAAATCAATGATGAT 2234

RESULT 10

US-10-793-626-4389/c
; Sequence 4389, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4389
; LENGTH: 3373
; TYPE: DNA

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4389

Query Match 1.4%; Score 40; DB 1; Length 3373;
Best Local Similarity 56.8%; Pred. No. 2.1;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;

QY 3 GRAGATTTCCARCACTTCAAGCACATATGATCCGTGTGGGTGGTAAATTCRAC 62
DB 2864 GAGGATTTATACAGCTTTGAAAATAATTAATTAATTAATTAATTAATTAATTA 2805
QY 63 TTGTAATRCATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 120
DB 2804 TTGCAATTAATTAATTAATTTTGTACTTAATTAATTAATTAATTAATTAATTA 2747

RESULT 11

US-10-793-626-3926/c
; Sequence 3926, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3926
; LENGTH: 3927
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3926

Query Match 1.4%; Score 40; DB 1; Length 3927;
Best Local Similarity 56.8%; Pred. No. 2.3;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;

QY 3 GRAGATTTCCARCACTTCAAGCACATATGATCCGTGTGGGTGGTAAATTCRAC 62
DB 1836 GAGGATTTATACAGCTTTGAAAATAATTAATTAATTAATTAATTAATTAATTA 1777
QY 63 TTGTAATRCATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 120
DB 1776 TTGCAATTAATTAATTAATTTTGTACTTAATTAATTAATTAATTAATTAATTA 1719

RESULT 12

US-10-793-626-4109/c
; Sequence 4109, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4109
; LENGTH: 3923
; TYPE: DNA

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4109

Query Match      1.4%; Score 39.8; DB 1; Length 3923;
Best Local Similarity 55.6%; Pred. No. 2.6;
Matches 74; Conservative 1; Mismatches 58; Indels 0; Gaps 0;

QY 53 AAATTCRCATCTGTTAATCTCTATTATTAATTTTAAATAATTAATCTATTATATCATATATAA 112
Db 2103 AATGATGACATTTTATAGTATATGACTATTATTTTAAATAAATTAACACATCCCTATTATA 2044

QY 113 TAATATTATTACTTATTAATAAAGTTAATAAAGTAACGAAATTAGGACTATAATCCC 172
Db 2043 CTGTACTACAGTCAAGATAGAGTGCGGATTAATAGTACTTGAATAATTACTATAATTAC 1984

QY 173 ATTGCCCTTAAAA 185
Db 1983 ACAAGTATTATA 1971

RESULT 13
US-10-793-626-2361
; Sequence 2361, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2361
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2361

Query Match      1.3%; Score 37.8; DB 1; Length 1809;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 72; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 23 AGCACATATTGATCCTCTGTGTGGTGGTAAATTCRCACCTGTATATCTTATTATTAAT 82
Db 322 AGCTCTGCATGGAGTTTTTTAGTACCTGAAATTTTACAACCTAGATGAAGCTACTATTCAA 381

QY 83 TTTTAAATTAATCTATTATCATATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 142
Db 382 TCTTTATCGAATCTAATGATGATTTAAACGATGATGATTCGATTTGAAATGATTAAT 441

QY 143 AAAAAGTAACGAAAT 157
Db 442 GAGAAAGCTCCACAT 456

RESULT 14
US-10-793-626-4033/c
; Sequence 4033, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04

```

```

; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4033
; LENGTH: 3591
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4033

Query Match      1.3%; Score 37.8; DB 1; Length 3591;
Best Local Similarity 53.3%; Pred. No. 7.5;
Matches 72; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 23 AGCACATATTGATCCTCTGTGTGGTGGTAAATTCRCACCTGTATATCTTATTATTAAT 82
Db 2984 AGCTCTGCATGGAGTTTTTTAGTACCTGAAATTTTACAACCTAGATGAAGCTACTATTCAA 2925

QY 83 TTTTAAATTAATCTATTATCATATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 142
Db 2924 TCTTTATCGAATCTAATGATGATTTAAACGATGATGATTCGATTTGAAATGATTAAT 2865

QY 143 AAAAAGTAACGAAAT 157
Db 2864 GAGAAACGCTCCACAT 2850

RESULT 15
US-10-793-626-4132/c
; Sequence 4132, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4132
; LENGTH: 2988
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4132

Query Match      1.3%; Score 37.2; DB 1; Length 2988;
Best Local Similarity 52.7%; Pred. No. 9.5;
Matches 78; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 64 TGTAAATRCATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 123
Db 509 TTTATATGTTAAAAAGTATTTCTCTTTTAAAAAAGTAAATAAATTTTAAAAAAGTAAATA 450

QY 124 CTTATATTAAAAAGTTAATAAAGTAACGAAATTAGGACTATAATCCCATTCCTTTAA 183
Db 449 CTAAAAATCAATTTTATACAATCCCTAAAAAGTCTATTATTTTGTGCAATTTGACTATA 390

QY 184 AATTAAACAAAGGAGTAATAGTGAAA 211
Db 389 TACTTGTCATTAATGCAAAATATATATA 362

Search completed: November 29, 2005, 08:13:49
Job time : 569 secs

```